

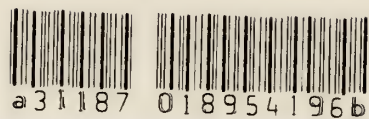
HC 115.B56X 1989

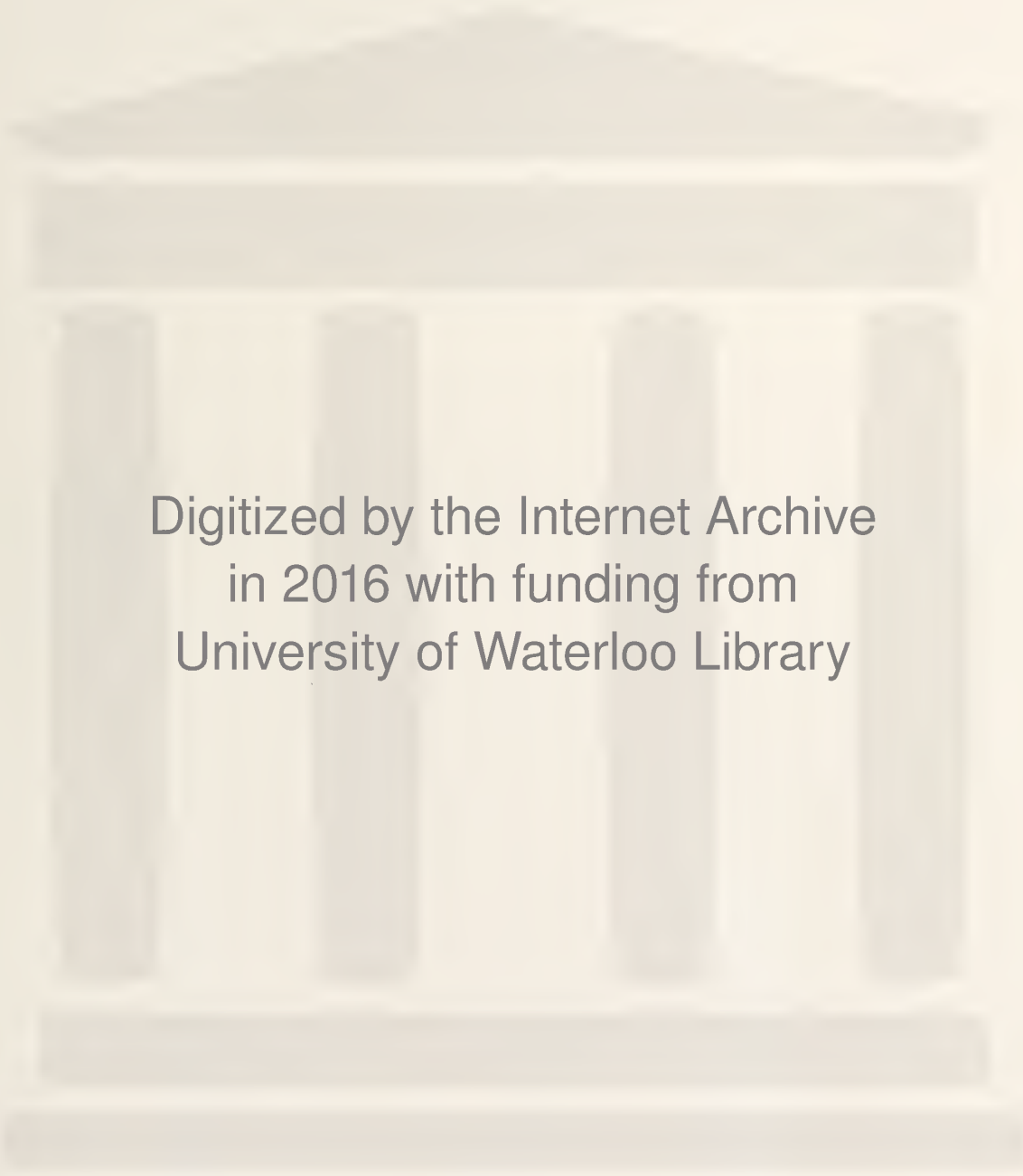


a31187 018954196b

HC
115
B56x
1989
Porter

HC 115.B56X 1989





Digitized by the Internet Archive
in 2016 with funding from
University of Waterloo Library

<https://archive.org/details/creatingcanind7100bloo>

CANADIAN INDUSTRY IN 1871

Research Report 4

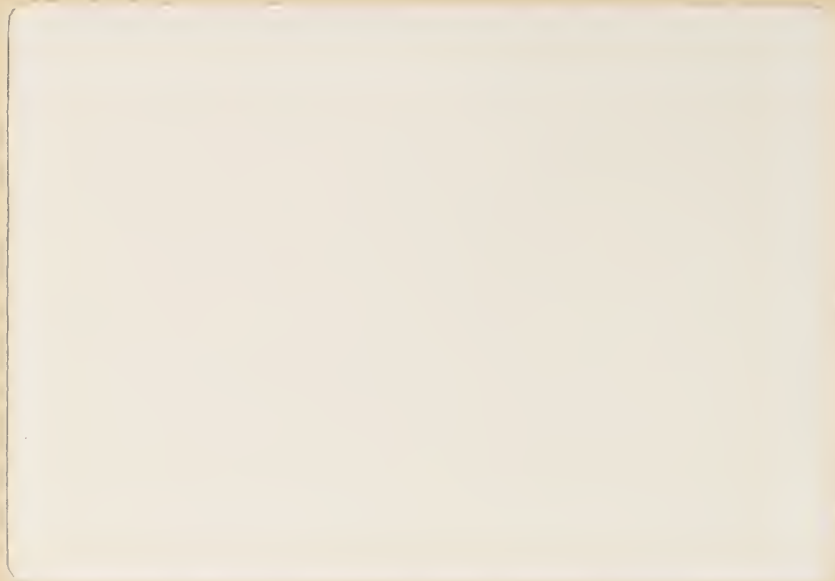
CREATING CANIND71: PROCEDURES FOR MAKING THE 1871 CENSUS MACHINE-READABLE

Elizabeth Bloomfield and G.T. Bloomfield

Elizabeth Bloomfield, series editor

November 1989





Cover Illustration (selected and described by G.T. Bloomfield)

The view of Whitevale, Township of Pickering, was chosen to illustrate Canadian industry in 1871. At this time, a high proportion of manufacturing activity was still located in small settlements, some of which were growing rapidly into towns. **Lovell's Directory** (1871) described Whitevale as:

A thriving village...[with] extensive flouring and woollen mills... Montreal Telegraph Co has an office here. Distant from Whitby, the county town, and a station of the Grand Trunk Railway, 13 miles. Mail daily. Population about 250.

Truman P. White has acquired the water rights at Majorville on Duffin's Creek in 1845 and developed a grist mill, a saw mill and, later, a woollen mill. By 1871 the census enumerated six significant industrial establishments employing 66 workers and with a total value of production amounting to \$125,000. The transition from waterwheels (70 horsepower) to steam engines (66 horsepower) was already apparent in the village by this date. In common with its counterparts across the country, Whitevale's basic industrial activities were closely associated with the local agricultural area. There was also considerable economic integration apparent in the ownership of several establishments by Truman P. White and in the making of staves in the sawmill for the cooper shop which in turn supplied the flour mill with basic containers for transporting the flour to market.

Unlike many of its contemporaries, Whitevale has remained about the same size ever since 1871. The 1971 census recorded a population of only 273 in the unincorporated settlement. Whitevale never achieved connection by railway, county road or provincial highway. Much of the surrounding land was acquired for the planned Pickering airport and new town in 1972/3 and today the settlement is threatened by the creation of a municipal solid waste dump for Metropolitan Toronto and the Durham Region.

The illustration was first published in the **Illustrated Historical Atlas of the County of Ontario** (Toronto: J.H. Bees and Co., 1877), reprinted Ross Cumming, 1972).

Research Report 4

**CREATING CANIND71:
PROCEDURES FOR MAKING THE 1871 CENSUS
MACHINE-READABLE**

Elizabeth Bloomfield and G.T. Bloomfield

Elizabeth Bloomfield, series editor

November 1989

"HC
115
B56x
1989
Porter

Department of Geography

University of Guelph

1989

Copies may be ordered from:

**CANADIAN INDUSTRY IN 1871 PROJECT
Department of Geography
University of Guelph
GUELPH, ONTARIO
N1G 2W1**

Copyright C 1989 Elizabeth Bloomfield

Canadian Cataloguing in Publication Data

Bloomfield, Elizabeth.
Creating CANIND71

(Canadian industry in 1871 ; 4)
Includes bibliographical references.
ISBN 0-88955-192-8

1. CANIND71 (Information retrieval system).
2. Canada - Industries - History - Data processing.
3. Canada - Census, 1871. I. Bloomfield, G. T. (Gerald T.). II. University of Guelph. Dept. of Geography. III. Title. IV. Series.

HC115.B58 1989 025.06'3380971 C89-090699-8

CANADIAN INDUSTRY IN 1871 (CANIND71) PROJECT

Between 1982 and 1989, a project based in the Department of Geography, University of Guelph, has made machine-readable the full data for some 45,000 industrial firms that were enumerated in Canada's first national census in 1871. A uniquely valuable source is thus becoming accessible to scholars and researchers in several disciplines. The 1871 schedules contain a wealth of information which was not published at the time or later. Although similar details were collected in the censuses of 1881, 1891, 1901 and 1911, none of the manuscript schedules for those years have survived. The CANIND71 database has great significance in being a detailed 'snapshot' of industrial activity just after Confederation, at a time of transition in industrial technology, business organization and work discipline. The records include examples of all kinds of industrial work environments from mills and artisanal craftshops in mainly rural settings to factories, manufactories and sweatshops in the growing towns and cities.

The CANIND71 project is important for its methodological experience in handling large quantities of historical data and making them accessible to users. Relevant aspects include the total coverage of all establishments and all variables recorded in the original source and our dedication to making the material available to others in a variety of software environments and with full explanation of the source and methodology. As well as the data for each establishment, we have added precise geographical references and Standard Industrial Classification codes (SIC) for all establishments, which permit both the retrieval of details for individual businesses and their systematic aggregation by industry type or geographical area.

Creation of the CANIND71 database has been assisted by several grants from the Social Sciences and Humanities Research Council of Canada between 1985 and 1989. The most substantial of these were Grants 482-87-0010 and 482-88-0010 to Elizabeth Bloomfield as principal investigator, in the Strategic Grants Program: Women and Work Theme. These grants, totalling \$114,000, supported the most intensive phase of database creation in 1988 and 1989. Other SSHRC grants to Elizabeth Bloomfield (principal) in 1985 and to Kris Inwood (principal) in 1988 have also helped. In addition, smaller grants from the University of Guelph to Gerald Bloomfield and Kris Inwood have supported the project for short periods. Personal funds have also been necessary. Some preliminary activity on the Maritime data during 1986 was helped by a grant from St Mary's University, Halifax, to Professor Inwood and Professor John Chamard.

The officials responsible for the original 1871 Census of Canada believed that the information they collected and collated was 'as accurate as is humanly possible.' In our turn, we are devoting several months in 1989-1990 to rigorously checking and editing the SAS datasets for Ontario, the Maritimes and Quebec on the mainframe computer. We expect that the final version of the whole database will be available for others to use from January 1991. Those interested in obtaining the whole database or partial datasets should contact Dr Elizabeth Bloomfield, C/- Department of Geography, University of Guelph, Guelph, Ontario, N1G 2W1, after September 1990.

CANADIAN INDUSTRY IN 1871 PROJECT: RESEARCH REPORTS

- 1. Industry in Ontario Urban Centres, 1870: Accessing the Manuscript Census,** Elizabeth Bloomfield, Gerald Bloomfield, Janine Grant and Peter McCaskell (1986).
- 2. Water Wheels and Steam Engines: Powered Establishments in Ontario,** Gerald Bloomfield and Elizabeth Bloomfield (1989).
- 3. The Ontario Urban System at the Onset of the Industrial Era, 1871,** Elizabeth Bloomfield and Gerald Bloomfield (1989).
- 4. Creating CANIND71: Procedures for Making the 1871 Industrial Census Machine-Readable,** Elizabeth Bloomfield and Gerald Bloomfield (1989).
- 5. Glossary of Industrial Language,** Jane Turner, Janine Grant and Barbara Sibley (1989).
- 6. French-English Dictionary of Industrial Language,** Jane Turner, Janine Grant and Barbara Sibley (1989).
- 7. Standard Industrial Classifications Applied to Historical Data: the Case of the 1871 Industrial Census,** Gerald Bloomfield and Elizabeth Bloomfield (1989).
- 8. Industrial Leaders: The Largest Manufacturing Firms in Ontario, 1871,** Elizabeth Bloomfield and Gerald Bloomfield (1989).
- 9. The Hum of Industry: Millers, Manufacturers and Artisans of Wellington County,** Elizabeth Bloomfield and Gerald Bloomfield (1989).

ACKNOWLEDGEMENTS

The interest and support of all who have assisted with this project are gratefully acknowledged. During the earlier phases, Janine Grant and Stephen Bellinger coded data for Ontario's urban places. From May 1985 to June 1989, Janine Grant was on the project staff, joined during the final 18 months by Barbara Sibley. The quality of the final database owes much to their careful and thorough work. Peter McCaskell, first as programmer-analyst in the Department of Geography and then from Computing Services, has helped substantially with database management and programming through all phases of the project. We appreciate the shelter provided to this project by the Department of Geography, University of Guelph throughout the 1980s. We are also grateful to the Social Sciences and Humanities Research Council of Canada: Strategic Grants Program for financial assistance during 1988 and 1989 which has enabled us to complete the creation of the CANIND71 database.

CONTENTS

	page
1 Significance of the census manuscript source	1
<i>Figure 1: Facsimile of English-language schedule</i>	6
<i>Figure 2: Facsimile of French-language schedule</i>	7
2 Project goals and methodology	11
Question of sampling versus recompiling the census	11
Transcription in natural language	12
Systems of locational references and industrial classification	13
<i>Figure 3: Geographical units for which 1871 census manuscript data may be aggregated and presented</i>	14
<i>Figure 4: Potential aggregations of 1871 census manuscript data</i>	16
<i>Table 1: Standard Industrial Classification: major groups</i>	17
Efficiency and accuracy of data entry and editing procedures	18
<i>Figure 5: Chronology of creating the CANIND71 database</i>	19
Accessibility in various hardware and software environments	22
3 Data structure	22
Basic variables	23
Derived variables	24
<i>Figure 5: Sample CANIND71 record: John Watson of Ayr</i>	26
<i>Table 2: Industrial establishments in Ayr district by proprietor</i>	27
<i>Table 3: Industrial establishments in Ayr district by employees</i>	27
<i>Table 4: Industrial establishments in Ayr district by output</i>	28
<i>Table 5: Industrial establishments in Ayr district by SIC code</i>	28
<i>Table 6: Summary of industrial data, North Dumfries, by CED</i>	29
<i>Table 7: Summary of industrial data, Waterloo South, by CED</i>	29
<i>Table 8: Summary of industrial data, Midwestern Ontario, by CED</i>	29
<i>Table 9: Industrial structure of North Dumfries by SIC code</i>	30
<i>Table 10: Industrial structure of North Dumfries by SEC</i>	31
<i>Table 11: Ontario: industry classes in the machinery sector by SIC</i>	32
<i>Table 12: Canada: machinery manufacturers with output over \$50,000</i>	33
4 Coding and editing procedures and problems	34
Proprior	35
Typeest	36
Typepow and Force	38
Kinds, quantities, values of Raw Materials and Products	38
Units of measurement	41
Comments	42
Industrial code	42

5	How good are the 1871 manuscript census data?	43
	Missing data	44
	Table 13: <i>Missing data in the CANIND71 database</i>	45
	Table 14: <i>Raw material and product data in two Quebec districts</i>	47
	Missing records	48
	Differences between published and manuscript census totals	48
	Table 15: <i>Variance between published and manuscript data 1</i>	49
	Table 16: <i>Variance between published and manuscript data 2</i>	50
Appendix A-1:	Census districts in 1871	53
Appendix A-2:	Standard Industrial Classification adapted for 1871: short list of types with at least 20 occurrences	55
Appendix A-3:	CANIND71 database: alphabetical list of variables	60
Appendix A-4:	Units of measurement and abbreviations	62

Among the new sources becoming available for Canadian historical research in several disciplines is the CANIND71 database that has been derived from the manuscript schedules of industrial establishments recorded in the 1871 Census of Canada. The procedures used to create the database during the past five years are described in this research report, as a guide to prospective users.

1 SIGNIFICANCE OF THE CENSUS MANUSCRIPT SOURCE

The manuscript schedules on industrial establishments, made available in the later 1970s on microfilm by the National Archives as part of the whole 1871 manuscript census, constitute a uniquely valuable source for Canada.¹ While the 1851 and 1861 census manuscripts are extant, their format is much more awkward to use, as the more limited industrial details are scattered through the household schedules.² The 1871 census was the first taken on a consistent basis for the provinces that formed the Dominion of Canada by that date. Its procedures were very carefully planned and implemented "with the utmost accuracy possible" by "honest, intelligent, well-instructed and pains-taking staff".³ Although similar details were collected in the censuses of 1881, 1891, 1901 and 1911, the manuscript schedules for industrial establishments for those years have not survived. Moreover, the 1871 schedules contain a wealth of information which was never published at the time. Not only are the 1871 census manuscripts the only systematic source surviving in such detail from the nineteenth century. They also mark a time of transition in industrial technology, business organization and work discipline.

How does this unique source for Canada compare with contemporary census records for other countries? In enumerating and compiling industrial information in 1871, Canada no doubt drew upon American experience, as the United States Census had gathered data on manufacturers since 1810. By 1870, the United States was not only publishing an 843-page volume on the "Wealth and Industry of the United States" but a statistical atlas as well.⁴ Some manuscript schedules have survived from the 1850, 1860, 1870 and 1880 censuses of industry in the United States. Samples of these data have been partially exploited during the past 15 years by the Bateman, Weiss and

¹ T.A. Hillman, Catalogue of Census Returns on Microfilm, 1666-1881 (Ottawa: Public Archives of Canada, 1981).

² A.A. Brookes, "Doing the Best I Can: The Taking of the 1861 New Brunswick Census", Histoire sociale IX (November 1976): 73-77.

³ The careful planning is evident in the manual prepared by the Minister of Agriculture for commissioners and enumerators charged with conducting the census: "Manual Containing the Census Act and Instructions to Officers Employed in the Taking of the First Census of Canada, 1871: An Act respecting the First Census (assented to 12th May, 1870)" in the Canada Sessional Papers No. 64 (1871). The quotation about the "utmost accuracy" is from page 120.

⁴ U.S. Bureau of the Census, Bureau of the Census: Catalog of Publications, 1790-1972 (Washington, DC, 1974).

Atack team at Indiana University, whose purposes and methodology may be compared with our own.⁵

In the United Kingdom, while the General Record Office established a high reputation with the 1851 Census collection of publications, the Census Act specifically excluded an industrial census. A partial tabulation of "masters" (employers) and "men" (employees) was compiled from the 1851 and 1861 occupational returns. It was recognized as being imperfect; "...it can only be rendered complete in the event of the Census being extended to an Inquiry into the Industry of the Country".⁶ Britain began systematically to collect industrial statistics as late as 1907, when it had lost its title of "workshop of the world". Only after World War II did Britain establish a regular series of industrial censuses with full publication of the results. None of the early manuscript returns for industrial establishments have survived.⁷ Some of the published data for nine Scottish burghs in the 1851 Census were used in an assessment of the characteristics of mid-nineteenth century industry.⁸

In Australia, as Linge has explained in his monumental study⁹, systematic collection of industrial statistics in a comprehensive manner began first in Victoria in 1868 and spread gradually to the other colonies. Western Australia established a full series in 1897. As in Britain, the initial impetus to gathering statistics on industry was the need for factory inspection. Comprehensive surveys to gather information on the value of production, the cost of inputs and value of investment came much later.

⁵ The U.S. data source and the Indiana University project procedures are discussed in Fred Bateman and Thomas Weiss, A Deplorable Scarcity: The Failure of Industrialization in the Slave Economy (Chapel Hill: University of North Carolina Press, 1981): 23-26 and Appendix A; Jeremy Atack, Estimates of Economies of Scale in Nineteenth-Century United States Manufacturing (New York and London: Garland Publishing, 1985): 40-81. The Philadelphia Social History Project has used all the manuscript industrial data for that city from the 1850 through 1880 censuses, as reported in Henry Williams, "Data Description", in Philadelphia: Work, Space, Family and Group Experience in the 19th Century, edited by Theodore Hershberg (New York: Oxford, 1981): Appendix II.

⁶ United Kingdom, Census 1861, General Report, p. 29. Republished in Irish University Press, British Parliamentary Papers Series, Population, vol. 15, p. 53.

⁷ Business Statistics Office, Historical Record of the Census of Production, 1907-1970 (London: HMSO, 1978).

⁸ R. Rodger, "Concentration and Fragmentation: Capital, Labor and the Structure of Mid-Victorian Scottish Industry", Journal of Urban History 14, 2 (1988): 178-213.

⁹ G.J.R. Linge, Industrial Awakening: A Geography of Australian Manufacturing, 1788-1890 (Canberra: Australian National University Press, 1979), Appendix 1.

New Zealand added a industrial schedule to its quinquennial census in 1867.¹⁰ It followed a very similar pattern to Canada until the early 1920s when the industrial census was separated from the census of population. All manuscript material prior to 1966 has been destroyed.

Data for industrial activity in the four provinces of Canada during the twelve months ending 31 March 1871 were recorded on the sixth of nine schedules used in taking the 1871 Census of Canada. The nine schedules were as follows: (1) nominal return of the living, including details of names, gender, ages, family relationships and dwellings; (2) nominal return of the deaths during the past twelve months; (3) return of public institutions, real estate, vehicles and implements; (4) return of cultivated land, of field products and of plants and fruits; (5) return of livestock, animal products, homemade fabrics and furs; (6) return of industrial establishments; (7) return of products of the forest; (8) return of shipping and fisheries; and (9) return of mineral products. As page and line references to the nominal returns had to be specified on each of Schedules 3, 4, 5, 7, and 8, it is possible now to interrelate the information in all these returns. Such cross-references are not possible with the industrial schedules, though enumerators were charged with asking every family they visited whether any member carried on any "Industrial Establishment".¹¹ The industrial schedule contained the only questions relating to dollar values in the whole census.

It is important to note that the census enumerators were instructed to record "all industry of any importance which is conducted in separate establishments or workshops." An industrial establishment was defined as "a place where one or several people are employed in manufacturing, altering, making up or changing from one shape into another, materials for sale, use or consumption, quite irrespectively of the amount of capital employed or of the products turned out." Examples of industrial establishments given in the "Manual" were "a lime kiln, a cheese factory, a brick-yard, a ship-building year", a grinding-stone factory, a sulphuric acid manufactory, a saw-mill, a marble cutter's shed, a wheel factory, a pottery, a foundry, a meat-curing or "packing" establishment, an establishment to manufacture copper regulus, or purify plumbago, a cloth manufactory, a carding mill, a grist mill, a planning and dovetailing mill, a sash factory, .. a shoe-making, harness-making, dress-making, tailor's or blacksmith's shop, or carpenter's or joiner's shop". It was stated that "all repairs, mending or custom work are understood to be industrial products; and are to be entered accordingly, by value, on the returns of industrial establishments".¹² Thus the definition of industrial activity was considerably broader than it would be in the twentieth century.

¹⁰ G.T. Bloomfield, New Zealand: A Handbook of Historical Statistics (Boston: G.K. Hall, 1984), p. 157.

¹¹ "Manual: Instructions to Officers", Sessional Papers (1871): 131.

¹² "Manual: Directions concerning the separate schedules", Canada Sessional Papers (1871): 138-9.

Enumerators were specifically instructed to record returns of Industrial Establishments in the geographical units - Division, Sub-District or Division-in which they were found "and nowhere else. The principle is essential in every case. The production is attached to the locality".¹³

Some additional instructions are useful in interpreting the census manuscript data. The values to be stated for raw materials and output were not to be confused with any concept of profit, a provision designed for custom work such as tailoring or work done on toll such as that of some grist mills or carding mills:

It matters not whether the raw material is in the ownership of the manufacturer or not, whether it is transformed on account of one or another person, whether the working is profitable or losing business; the information required is the result from the establishment, so far as requested. For instance, a saw mill may saw logs and other lumber for a great number of other persons; the working may cost more than the returns bring to the owner; but, nevertheless, the amount of raw material has changed form, and so much value has been added to it; and this is the fact to be recorded.¹⁴

Numbers of workers or "employees" were to include only those persons actually working in the industrial establishment:

...The number of people employed may be made up exclusively with members of the family of the proprietor; in other places the proprietor and family may not form part of the people employed.¹⁵

Though enumerators were invited to specify types, quantities and values of individual raw materials and products of each establishment, it was foreseen that usually only the aggregate values would be stated:

In many cases the raw materials, or articles manufactured are of such a multifarious character that they must be lumped together, and entered by the value.¹⁶

Some other features of the 1871 industrial census may be noted, though they were not explicitly prescribed in the "Manual". No minimum value of output was set, in contrast to the United States, where only establishments with at least \$500 worth were included. The values of fixed capital and floating capital invested in the business were distinguished rather than merged and the number of working months was specified. Although the industrial workforce was subdivided by age and gender into men, women, boys (under 16

¹³ Ibid.

¹⁴ Ibid.

¹⁵ Ibid.

¹⁶ Ibid.

years) and girls (under 16 years), only "average numbers" of each employed during the census years were to be entered and the amount of wages paid to each group of worker was not separately stated.¹⁷

Enumerators were not specifically instructed on how to determine whether female workers should be recorded on the industrial schedules. But their recording practices were probably influenced by the directions for entering details of occupation on the nominal schedules. For that schedule, census enumerators were instructed, that "in the case of women, unless they have a definite occupation besides their share in the work of the family or household", they were to be entered as having no "occupation".¹⁸ Yet it was apparently expected that men who were primarily farmers but also produced shingles, sawn lumber, lime or potash for a few weeks in the year would have these industrial activities recorded on Schedule 6. It seems likely that enumerators may have been deterred by this strict requirement from entering small industrial enterprises that were carried on part-time by women.

As already noted, all repairs, mending and custom work done by industrial establishments were to be included. However, a little ambiguously, the 1871 industrial census schedules did not include "products of domestic industries such as building, furniture making, clothing, tools, boat and carriage building" by farmers and the "seaside populations." Such domestic industry was believed to be considerable, especially in the Maritimes and Quebec. For example, the quantity of woolen and linen cloth produced by the families of farmers was estimated to be roughly equal to that produced by all the large manufacturers.¹⁹ It would appear that enumerators did not all apply any distinction between domestic and general industrial activity in the same way. Some of them entered details for such activities as weaving and clothing on the industrial schedules which were later apparently not included in the published tables. Others seemed not to record such activities on the industrial schedules and presumably did so on the agricultural schedules.

The 1871 census enumerators recorded the following details for each establishment they found, relating to industrial operations in the twelve months preceding 1st April, 1871. These details were handwritten in numbered columns horizontally across the schedule forms, five establishments to the page. Figures 1 and 2 illustrate forms in both English and French languages.

¹⁷ The range of variables is similar to that in the U.S. censuses for 1850-1870, except that the 1871 Census of Canada recorded "floating" or working capital in addition to "fixed" capital (real and personal estate), the number of working months in the year, and the number of girls employed as well as of men, women and boys. On the other hand, the Canadian data do not permit a separation of the wage costs of men, women and youths.

¹⁸ "Manual: Directions concerning Schedule 1" Canada Sessional Papers (1871): 134.

¹⁹ Census of Canada 1871, Volume 3, p. viii.

Ontario
District No. 77
Sub-District 2
Schedule No. 6.—Return of Industrial Establishments.

Province of
Census of 1871.

Victoria Ward
for 2 of 6

Name of Industrial Establishment, and other such information.	Kind of Capital Employed, in \$.	Number of working months in the year.	Average Number of People employed.				Moving Power.		Raw Material.			Products.		
			Over 10 Years.		Under 10 Years.		Kind.	Horse Power.	Kind.	Quantities.	Aggregate Value, in \$.	Kind.	Quantities.	Aggregate Value, in \$.
			Male.	Female.	Boys.	Girls.								
1. Banding, Tullies & Drapery Mill, Tullies A.	10,000	1000	12	3	—	—	Wate	20	Wool	30,000 lbs	7500	Wool	20,000	9000
2. The Factory, Washington Sexton.	4000	4000	12	12	—	—	Wate	100	Wool	40	2400	Wool	10,000	3480
3. Saw Mill, Penley & Pattee.	150,000	300,000	12	250	—	—	Wate	500	Saps	150,000	240,000	Wool	10,000	15,000
4. Flour Mill, Young Brown & Co.	20,000	6000	12	6	—	—	Wate	100	Wheat	100,000	10,000	Wheat	20,000	120,000
5. Flour Mill, Morrison & Co.	30,000	15,000	12	12	—	—	Wate	100	Wheat	100,000	10,000	Wheat	20,000	120,000

Figure 1: Facsimile of English-language schedule

Province of Ontario		District No. 24 United Ottawa		Sous-District St. Lawrence Reg.		Page 1		Recensement de 1871		
Recensement de 1871		Tableau No. 9 — Etablissements Industriels		Division No. 2, St. Lawrence		Recensement de 1871		Page 1		
Noms des propriétaires ou des occupants de l'usine	Capital en dollars	Capital en francs	Noms des usines	Natures	Produit principal	Moyenne du nombre de personnes occupées		Moyens de transport	Matières premières	
						As de 16 ans et au-dessous de 16 ans	Plus de 16 ans			
						Hommes	Femmes			
						1	2			
						3	4			
						5	6			
						7	8			
						9	10			
						11	12			
						13	14			
						15	16			
						17	18			
						19	20			
						21	22			
						23	24			
						25	26			
						27	28			
						29	30			
						31	32			
						33	34			
						35	36			
						37	38			
						39	40			
						41	42			
						43	44			
						45	46			
						47	48			
						49	50			
						51	52			
						53	54			
						55	56			
						57	58			
						59	60			
						61	62			
						63	64			
						65	66			
						67	68			
						69	70			
						71	72			
						73	74			
						75	76			
						77	78			
						79	80			
						81	82			
						83	84			
						85	86			
						87	88			
						89	90			
						91	92			
						93	94			
						95	96			
						97	98			
						99	100			
						101	102			
						103	104			
						105	106			
						107	108			
						109	110			
						111	112			
						113	114			
						115	116			
						117	118			
						119	120			
						121	122			
						123	124			
						125	126			
						127	128			
						129	130			
						131	132			
						133	134			
						135	136			
						137	138			
						139	140			
						141	142			
						143	144			
						145	146			
						147	148			
						149	150			
						151	152			
						153	154			
						155	156			
						157	158			
						159	160			
						161	162			
						163	164			
						165	166			
						167	168			
						169	170			
						171	172			
						173	174			
						175	176			
						177	178			
						179	180			
						181	182			
						183	184			
						185	186			
						187	188			
						189	190			
						191	192			
						193	194			
						195	196			
						197	198			
						199	200			
						201	202			
						203	204			
						205	206			
						207	208			
						209	210			
						211	212			
						213	214			
						215	216			
						217	218			
						219	220			
						221	222			
						223	224			
						225	226			
						227	228			
						229	230			
						231	232			
						233	234			
						235	236			
						237	238			
						239	240			
						241	242			
						243	244			
						245	246			
						247	248			
						249	250			
						251	252			
						253	254			
						255	256			
						257	258			
						259	260			
						261	262			
						263	264			
						265	266			
						267	268			
						269	270			
						271	272			
						273	274			
						275	276			
						277	278			
						279	280			
						281	282			
						283	284			
						285	286			
						287	288			
						289	290			
						291	292			
						293	294			
						295	296			
						297	298			
						299	300			
						301	302			
						303	304			
						305	306			
						307	308			
						309	310			
						311	312			
						313	314			
						315	316			
						317	318			
						319	320			
						321	322			
						323	324			
						325	326			
						327	328			
						329	330			
						331	332			
						333	334			
						335	336			
						337	338			
						339	340			
						341	342			
						343	344			
						345	346			
						347	348			
						349	350			
						351	352			
						353	354			
						355	356			
						357	358			
						359	360			
						361	362			
						363	364			
						365	366			
						367	368			
						369	370			
						371	372			
						373	374			
						375	376			
						377	378			
						379	380			
						381	382			
						383	384			
						385	386			
						387	388			
						389	390			
						391	392			
						393	394			
						395	396			
						397	398			
						399	400			
						401	402			
						403	404			
						405	406			
						407	408			
						409	410			
						411	412			
						413	414			
						415	416			
						417	418			
						419	420			
						421	422			
						423	424			
						425	426			
						427	428			
						429	430			
						431	432			
						433	434			
						435	436			
						437	438			
						439	440			
						441	442			
						443	444			
						445	446			
						447	448			
						449	450			
						451	452			
						453	454			
						455	456			
						457	458			
						459	460			
						461	462			
						463	464			
						465	466			
						467	468			
						469	470			
						471	472			
						473	474			
						475	476			
						477	478			
						479	480			
						481	482			
						483	484			
						485	486			
						487	488			
						489	490			
						491	492			
						493	494			
						495	496			
						497	498			
						499	500			
						501	502			
						503	504			
						505				

1. Kind of Industrial Establishment, Name of Proprietor or Company, and other such information / *Genres d'établissements industriels, noms de propriétaires ou des compagnies et autres renseignements.*

2. Fixed Capital invested in \$ / *Capital fixe, en piastres.*

3. Floating Capital employed in \$ / *Capital flottant, en piastres.*

4. Number of working months in the year / *Nombre de mois de travail dans l'année.*

5 - 8. Average Number of People employed / *Moyenne de nombre de personnes employées:*

5. Male over 16 years / *Hommes au-dessus de 16 ans;*

6. Female over 16 ans / *Femmes au-dessus de 16 ans;*

7. Boys under 16 years / *Garçons au-dessous de 16 ans;*

8. Girls under 16 years / *Filles au-dessous de 16 ans.*

9. Aggregate amount of Yearly Wages in \$ / *Montant collectif de gages durant l'année en piastres.*

10. Moving Power: Kind / *Force motrice: Nature.*

11. Moving Power: Nominal force / *Force motrice: Force nominale*
(stated in units of horse power)

12 - 14. Raw Materials / *Matières brutes:*

12. Kind / *Espèces;* (units of measurement often stated here as well)

13. Quantities / *Quantités;*

14. Aggregate Value in \$ / *Montant de la valeur en piastres;*

15 - 17. Products / *Produits:*

15 Kind / *Espèces;* (units of measurement often stated here as well)

16. Quantities / *Quantités;*

17. Aggregate Value in \$ / *Montant de la valeur en piastres;*

18. Remarks / *Remarques* (this space was sometimes used for additional information, perhaps an address, or an explanation of special difficulties encountered by the enumerator).

In addition to the data given by the census enumerator specifically for each industrial establishment in the prescribed columns, each schedule was headed with details of the relevant Province, Census District, Sub-district and Census Enumerator's Division, both the place names and the census number and letter codes. The census districts were intended to coincide with the districts defined for federal electoral purposes. In Ontario they consisted of aggregations of incorporated municipal units such as townships, towns and villages. Census sub-districts usually coincided with legally defined units such as townships, towns, villages or wards (in the larger urban centres) in Ontario. The census districts and sub-districts used in the other three provinces are

harder to compare with other territorial units.²⁰ Census sub-districts were usually subdivided into two or three tracts, for each of which an individual census enumerator was responsible. These tracts, known as divisions at the time and as CEDs (census enumerators' divisions) in this project, were not defined or mapped in any surviving records. Only with very careful field work, effectively retracing the steps of the census enumerator on his rounds, could one reconstruct the boundaries of these small units.

Only a very limited amount of this material was published in the official census volumes of the 1870s.²¹ The published statistics were organized primarily by various industrial types which were defined pragmatically rather than systematically and listed alphabetically. For each industry type, whether as specialized as whip making or as ubiquitous as blacksmithing, figures were published for numbers of establishments, hands employed, yearly wages, value of raw materials and value of products in each census district. Table 54 of Volume III summarized the aggregate data for all industrial activity in each census district under the headings: total capital invested, employees, wages, raw materials and value of products. Table 55 summarized by each province the aggregate data for each type of industrial activity.

Because of the primary organization of the published material by industrial type, it has been a major exercise to reconstitute the total industrial structure of particular census districts, a task which is now facilitated by computer methods. Summary data only were published for each of the 206 census districts defined in the four provinces, 90 of them in Ontario, 83 in Quebec, 14 in New Brunswick and 19 in Nova Scotia.

No industrial data at all were published for smaller areal units, either as summaries of total industrial activity or for specific types of industry. Thus the only 1871 industrial information published for urban centres was for the six cities, the boundaries of which coincided with those of one or more census districts. These were Montreal, Toronto, Hamilton, Ottawa, London and Kingston. No details were made available for individual establishments, and the material collected on the use of inanimate power was not released in any form. Furthermore, the published totals seem to have understated the real extent and values of industrial activity as these may be reconstituted from the manuscript census schedules.

The summary industrial data published in the 1871 census volumes have been used in important surveys of industrial activity in late nineteenth-century

²⁰ The 206 census districts and 1701 sub-districts were defined in Census of Canada 1871, Volume 5, pp. 338-454. Rather generalized maps of the census districts only have survived in the printed instructions and published census volumes. Otherwise we have only the legal descriptions of the census units, often in terms of other boundaries which may no longer have meaning. Reconstituting the territorial units used in the 1871 census has been a major exercise that is described in Boundaries of Canadian Census Units in 1871, #10 in this series of research reports.

²¹ Census of Canada 1871, volume III, Tables 28-55.

Canada, such as those by Chambers and Bertram and Gilmour.²² But these scholars had to make inferences based on totals for whole census districts (counties or portions of counties) and were limited by the implied definitions of industrial types employed by the 1871 census organizers.

Since the 1871 manuscript schedules were first released in the early 1970s, several scholars have examined them to support studies of particular cities, districts or industrial types.²³ However, these uses of the 1871 manuscript data have been specific to each researcher's limited purpose. Different systems of classification have been used, so that comparisons with other places or industrial sectors or with later periods are almost impossible. Very few of these users have transformed the data into machine-readable form, and those who have done so have been unconcerned about making their records accessible to other researchers. Scholars interested only in one category or sector of industrial activity may be daunted by the time-consuming labour of searching through the microfilmed schedules for a handful of establishments, and could easily miss some which were slightly misplaced in storage or the microfilming process.

²² E.J. Chambers and G.W. Bertram, "Urbanization and Manufacturing in Central Canada, 1870-1890," in S. Ostry and T.K. Rymes, eds. Papers on Regional Historical Statistics (Toronto: University of Toronto Press, 1966); J.M. Gilmour, Spatial Evolution of Manufacturing, Southern Ontario 1851-1891 (University of Toronto, Department of Geography Research Publications, 1972).

²³ Gregory Kealey, for example, used the Toronto data to provide a context for his study of industrial workers, in G.S. Kealey, Toronto Workers Respond to Industrial Capitalism (Toronto: University of Toronto Press, 1980). The York Social History project, directed by Michael Katz, coded data for industrial establishments in Hamilton, as part of its analysis of industrial capitalism in that city: M.B. Katz, M.J. Doucet, and M.J. Stern, The Social Organization of Early Industrial Capitalism (Cambridge: Harvard University Press, 1982). L.D. McCann has used the manuscript schedules for Halifax-Dartmouth and for Pictou County, together with Dun credit ratings: L.D. McCann, "The Mercantile-Industrial Transition in the Metals Towns of Pictou County, 1857-1931", Acadiensis 10, 2 (1981): 29-64. Eve Martel reconstituted the general patterns of industry in Montreal: E. Martel, "L'industrie à Montréal en 1871" (M.A. Thesis, Université du Québec à Montréal, 1977). Joanne Burgess has studied the organization of the Montreal shoe-making industry: J. Burgess, "L'industrie de la chaussure", Revue d'histoire de l'Amérique française 31 (1977): 187-210. Paul Craven and Tom Traves have drawn on census manuscript data for evidence of industrial activity in railway workshops and yards: P. Craven and T. Traves, "Canadian Railways as Manufacturers, 1850-1880", Canadian Historical Association Historical Papers (1983): 254-281. Jim Burant has featured the photographic studios of Saint John, and Ian McKay the confectionery and baking industry of Halifax: J. Burant, "A Written Portrait: Saint John Photographers and Their Studios in the 1871 Census", Archivaria 17 (1983-84): 275-7; I. McKay, "Capital and Labour in the Halifax Baking and Confectionery Industry During the Last Half of the Nineteenth Century", Labour/Le Travailleur 3 (1978): 63-108.

2 PROJECT GOALS AND METHODOLOGY

The project reported here has made machine-readable the manuscript data for over 45,000 industrial establishments in Canada in 1871. The project's methodology has been designed to make this information accessible, in systematic, standardized and machine-readable format, to serve the research interests of a wide variety of academic and applied historical research. It is expected that the database will interest historical and industrial geographers; economists and economic historians; business, labour, social and urban historians; industrial archaeologists and historians of material culture and technology. The experience of creating a large database from manuscript census material has methodological interest also for social scientists, quantitative historians and information specialists.

The main features of the CANIND71 methodology have been designed to facilitate maximum access to the data by prospective users. They may be summarized as follows:

- * Recording of data from all manuscript schedules, rather than only a sample of the industrial establishments counted in the 1871 census.
- * Transcription of all data from the schedules, in natural language rather than predetermined codes or a thesaurus of prescribed terms, and in the language, English or French, that was used by the original enumerator.
- * Addition to the basic records of precise geographical references and industry group codes in order to provide points of access to individual establishments as well as to aggregate data for places and for industry types.
- * Painstaking accuracy and consistency in data entry and systematic editing, so that the database would serve the needs of all kinds of users and the task of making the industrial schedules of the 1871 Census machine-readable would never have to be repeated.
- * Accessibility in various hardware and software environments.

We comment generally on various procedural problems and choices and then discuss more detailed procedures in compiling and using the database.

Question of sampling versus recompiling the census

Could we have used a sample rather than choosing "the monumental task of recompiling the census"? Sampling techniques are commonly used by social scientists in analyzing historical demographic data from the nominal census, though historians have usually chosen to capture the full population, at least for defined territorial areas.²⁴ In its use of the United States manuscript

²⁴ To quote some Ontario examples of the use of samples versus total coverage of a census source, the sociologists Darroch and Ornstein used a stratified, random sample of 10,000 households drawn from the 1871 nominal census of all four Canadian provinces: G. Darroch and M. Ornstein, "Family

censuses of manufacturing for 1850, 1860, 1870 and 1880, the Indiana University team took a random sample of about 200 firms, together with full details for the largest 20 firms in each state.²⁵ But not only would the U.S. project have had to consider formidable total numbers of establishments in several census years; its objectives were more limited than ours in serving only economic historians who were interested mainly in aggregate data. Factors in our decision for total coverage were the uniqueness of the source in Canada, the inadequacy of the published data as a universe for any sample, and the many and varied research purposes we expect to serve by making the manuscript material machine-readable.

We intend to serve a wider variety of research interests, including those concerned with the rich variety of individual enterprises in particular areas as well as the aggregate patterns. Scholars and researchers in some disciplines -- such as social and labour history and the history of material culture and technology -- may be primarily interested in retrieving data for individual establishments or for small industrial groups or geographical areas. For such users, the aggregate patterns of the database provide a comparative context. Geographers and economic historians place a greater emphasis on the aggregate patterns of industrial activity, which may now be drawn more accurately and also be illuminated by a broad variety of individual firm experiences. We decided against a sample because of our goal of linking macroscale and microscale, the broad generalizations with the particular details for individual firms, for industry groups, or for communities, cities and regions.²⁶

Transcription in the natural language of the original schedules

For the original pilot project on the 1871 manuscript industrial schedules in 1982, we were constrained by the available technology of 80-column cards

Co-Residence in Canada in 1871: Family Life-Cycles, Occupations and Networks of Mutual Aid", Historical Papers (1983): 30-55. Historians in the Canadian Social History Project (1967-73) on the city of Hamilton and in the Peel County History Project have opted for total coverage of the whole populations of their more limited geographical areas, as reported in M.B. Katz, The People of Hamilton, Canada West: Family and Class in a Mid-Nineteenth-Century City (Cambridge: Harvard University Press, 1975); D. Gagan, Hopeful Travellers: Families, Land and Social Change in Mid-Victorian Peel County, Canada West (Toronto: University of Toronto Press, 1981).

²⁵ The Indiana University project's sampling procedures yielded a sample size of 5,904 for 1850, 6,328 for 1860 and 4,859 for 1870, with sample proportions varying from .01 to 1.00 from state to state: Atack, Estimation of Economies of Scale, 44. The basic sample was supplemented with a dataset of the largest 20 firms in each state.

²⁶ For a general discussion of the relationships between the approaches of traditional history and social science history, relevant to this aspect of our methodology, see: J. Sharpless, "Collectivity, Hierarchy and Context: The Theoretical Framework for the Aggregation Problem", Historical Methods 17, 3 (1984): 132-140; I. Winchester, "History, Scientific History and Physics", Historical Methods 17, 3 (1984): 95-106.

keypunched for processing on the mainframe computer. We had to limit the numbers of variables to mainly numeric data and to use some abbreviated codes. But from 1986, as it became possible to use larger numbers of longer fields, we decided to transcribe all information found on the manuscript schedules and applied this rule retroactively to the records that had already been made machine-readable. Thus the database may be described in the main as a machine-readable "facsimile" of the original material, according to a scheme that avoids cryptic numeric or letter codes.

It was also decided, as much as possible, to follow the original language and expression of the manuscript information. Thus data were transcribed in French or in English, depending on the language used by the census enumerators. The original terminology and expression of the enumerators were retained rather than replaced with modern systems of terminology. There were two main reasons for this decision. In the first place, it was easier for project staff to transcribe consistently than to attempt simultaneously to decipher the manuscript information and translate and/or transform it into a controlled vocabulary. Moreover, we realized that the variant terms used by the enumerators could have intrinsic interest and meaning for our understanding of the organization of industrial activity in 1871.²⁷

However, the manuscript data were systematized in some cases, to allow for more orderly data retrieval. For example, the surname of the proprietor was consistently entered before the forename if the enumerator had not done this. Spelling errors were corrected and minor variations were standardized. Descriptions in French of the type of industrial activity usually began with a generic term such as "Moulin de..", "Atelier de..", "Boutique de ..", "Fabrique de.." before giving the specific type. We transposed the information, stating the specific type first as a keyword for sorting and retrieval purposes, then adding the generic term as an abbreviation. "Moulin de farine", for example, was entered in the computer record as "Farine, M", and "Boutique de forgeron" as "Forgeron, B". Practices of this sort, which were designed to increase the clarity and consistency of the database, are described more fully later in this report.

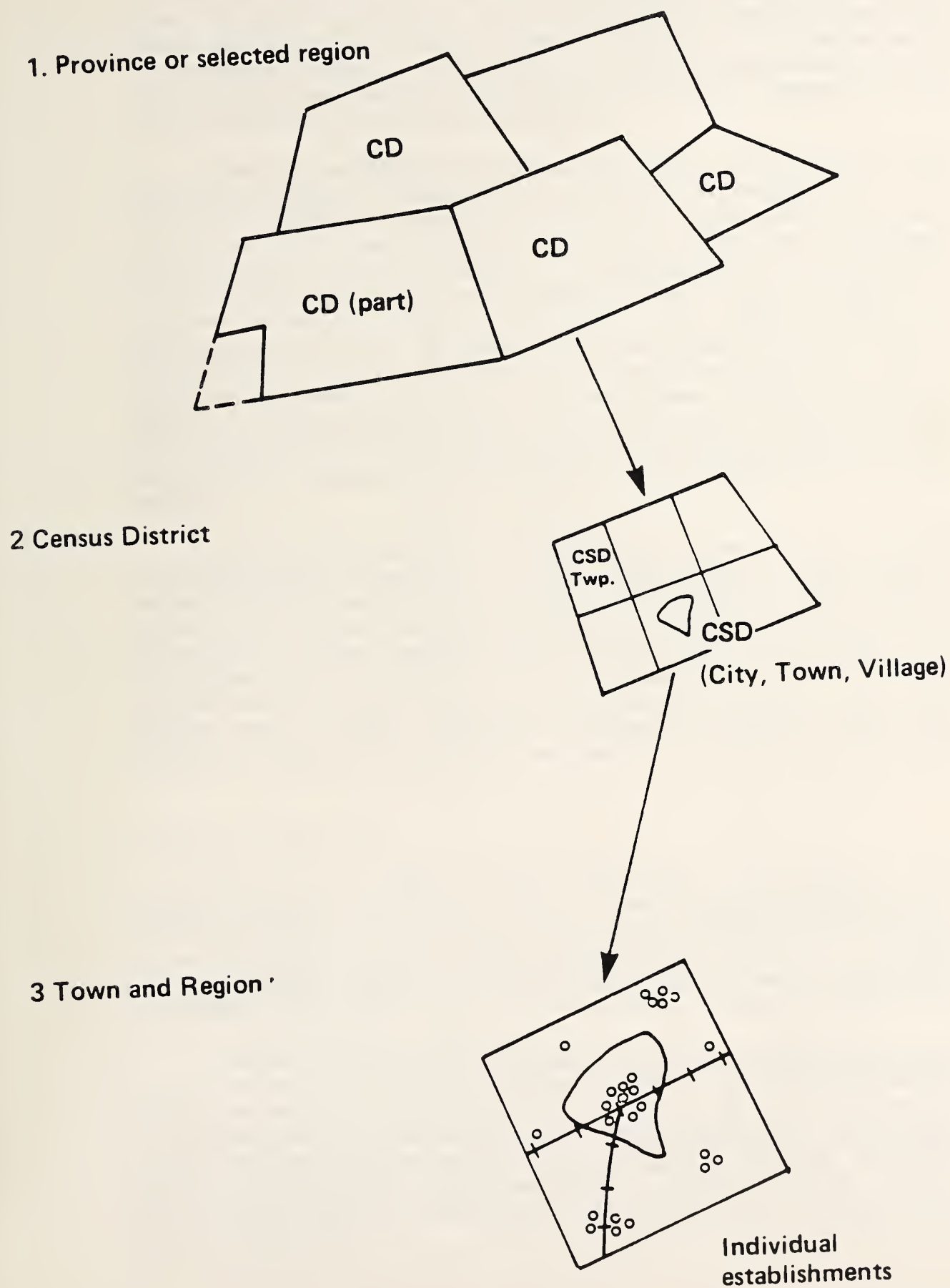
Systems of locational references and industrial classification

In order to be able to retrieve individual records and also to aggregate them by geographical area and industry type, specific locational and industry codes were added to each record.

Geographical details of town/township name, census district (CD) name and number, census sub-district (CSD) and census enumerator's division (CED) were coded in each case. Usually this information is provided in the marginal information on the manuscript schedules, though we also consulted the more

²⁷ For a guide to terminology used in the 1871 census manuscripts, see two reports in this series compiled by project staff, Jane Turner, Janine Grant and Barbara Sibley, Glossary of Industrial Language (#5) and French-English Dictionary of Industrial Language (#6).

Figure 3: Geographical units for which 1871 census manuscript data may be aggregated and presented



systematic sources.²⁸ For ease in returning to the original schedules, we followed the same numerical sequence as the census organizers, numbering census districts from 001: Essex in southwest Ontario to 206: Richmond, Nova Scotia, and using the same code letters and numbers for the smaller census sub-divisions. Figure 3 illustrates the hierarchy of geographical units for which the industrial data may be presented and aggregated. Appendix A-1 lists the codes and names of the 206 census districts.

A Standard Industrial Classification (SIC) code was added for each establishment record.²⁹ This system was adapted to the conditions of 1871 in two ways. Suffixes of hyphen and capital letter were added to the basic 3-digit code to give greater specificity, a brewery (109-B) being distinguished from a distillery (109-D), for example. An establishment which combined two or more products or services was designated with SIC codes of the two most important linked by a slash (/), so that a flour mill-cum-distillery is represented as 105/109-D. SIC codes were also generalized into Major Industry Groups, such as 5.01 for all Food and Beverage Industries; these are code-named SECs (Sectors) in the project and the database. Table 1 sets out the codes for the major industry groups (SECs) with the range of specific industry types (SICs) included in each, while Appendix A-2 lists the more common SIC codes that were assigned in making the 1871 industrial census data machine-readable.

Use of the SIC classification enables the user to cope with the full range of establishments enumerated in the 1871 census, when a wider range of industrial activities and processes was included in definitions of industry than would be today. These include agricultural services (SEC 1), forestry (SEC 2), fishing (SEC 3), mines, quarries, oil and salt wells (SEC 4), construction (SEC 6), utilities (SEC 7), trade (SEC 8), and services (SEC 10). By excluding establishments with these SIC codes, it is possible to create more select datasets from the original database for comparative purposes, limiting one's attention to activities which are now more narrowly defined as manufacturing

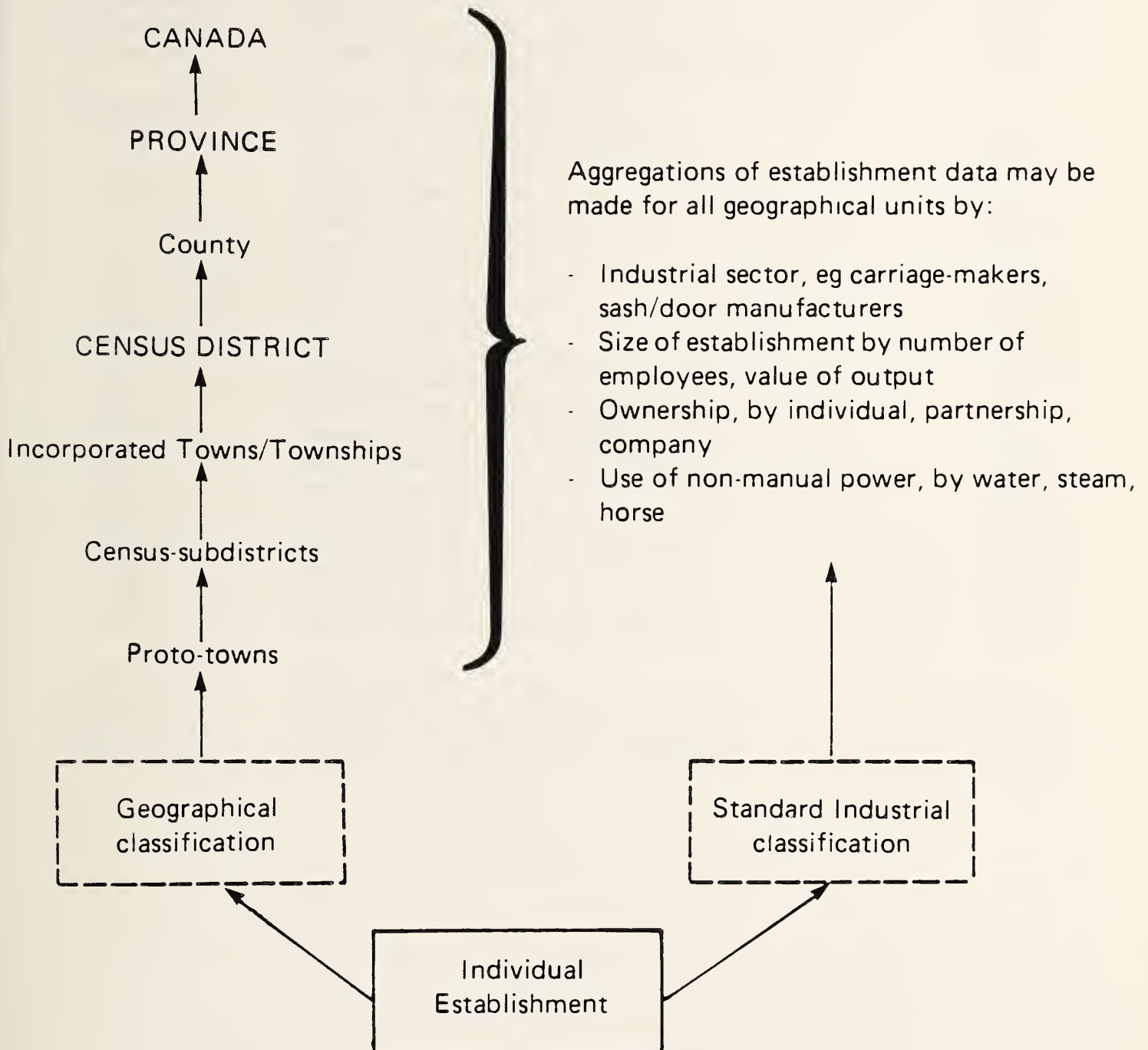
²⁸ A list of the census districts used in 1871 was appended to the "Manual" relating to the taking of the 1871 census, Canada Sessional Papers, 64 (1871): 142-150. In more detail, each census district was subdivided into census sub-districts, as specified in the Canada Gazette 4 (1870-71) and Census of Canada 1870-71, Volume 5, pp. 388-434. See also G.T. Bloomfield, Boundaries of Canadian Census Units in 1871, #10 in this series of research reports.

²⁹ Dominion Bureau of Statistics, Standard Industrial Classification Manual (revised edition, 1970). For a discussion of standard industrial classifications and of our reasons for choosing the 1970 system for use in this project, see G.T. Bloomfield and Elizabeth Bloomfield, Standard Industrial Classifications Applied to Historical Data: the Case of the 1871 Industrial Census, #7 in this series of research reports. The report includes also lists of all SIC codes in three systems of order: the logical order of the classification, by English name and by French name.

Figure 4: Potential aggregations of 1871 census manuscript data

Geographical Hierarchy

Characteristics of Industry



CAPITAL letters show the only geographical units for which data were published in 1871 Census

Table 1
Standard Industrial Classification: Major Groups

Division/Major Group		SIC codes
Division 1	Agricultural Services	001-029
Division 2	Forestry	031-039
Division 3	Fishing	041-049
Division 4	Mines, Quarries, Oil and Salt Wells	051-099
Division 5	Manufacturing Industries	
5.01	Food and Beverage Industries	101-109
5.02	Tobacco Products	151-159
5.03	Rubber Industries	161-169
5.04	Leather Industries	171-179
5.05	Textile Industries	181-189
5.06	Knitting Mills	231-239
5.07	Clothing Industries	241-249
5.08	Wood Industries	251-259
5.09	Furniture Industries	261-269
5.10	Paper Industries	271-279
5.11	Printing and Publishing	281-289
5.12	Primary Metal Industries	291-299
5.13	Metal Fabricating Industries	301-309
5.14	Machinery Industries	311-319
5.15	Transportation Equipment	321-329
5.17	Non-metallic Mineral Products	351-359
5.18	Petroleum and Coal Products	361-369
5.19	Chemical Industries	371-379
5.20	Miscellaneous Manufacturing	391-399
Division 6	Construction Industry	421
Division 7	Utilities, including gas and water	574-576
Division 8	Trade, including repair services	601-699
Division 10	Personal and Business Services	861-899

industry. Establishments may also be sorted as to whether they were primary (processing) or secondary (fabricating) industries.³⁰

The specification of locational references and industry classification in our project enables data for any single establishment to be retrieved easily. Details for groups of establishments may also be aggregated by geographical unit or industrial type. The records of industrial establishments are primarily retrievable in the order in which they were enumerated, within the framework of census district (CD), census sub-district (CSD) and enumerator's division (CED). But the records may be indexed on any field to achieve other systems of order - alphabetically by proprietor's name, by SIC code, or in size order (using value of production, numbers employed, or value of capital invested) for any geographical unit or grouping of units.

The entire database may also be indexed or sorted *en bloc*, without regard to location. It is possible to group all the establishments of a given industrial type by indexing on the SIC code, or to pick out the largest businesses by indexing on any numeric variable. Because individual establishment records are precisely coded for geographic location and industrial classification, their numeric data may be aggregated into successively larger groupings or areas. Figure 4 illustrates the possibilities of aggregating data for specific establishments geographically or by industry group.

Efficiency and accuracy of data entry and editing procedures

Any large project that seeks to capture historical census data will be labour-intensive. Whether volunteers or paid staff are used, clear procedures must be devised and followed in order to standardize the data and minimize costly errors, editing and corrections. It is necessary also to consider the ergonomics of the project, in relation to the stresses of working with manuscript records and video display terminals, including poor calligraphy and spelling, inferior microfilms and microfilm readers, and the cheaper computer monitors.

The beginnings of the project can be traced to the interest of Gerald and Elizabeth Bloomfield in the industrial development of urban places in the Grand River Valley of Ontario and to Elizabeth Bloomfield's postdoctoral research on urban elites and municipal inducements in Ontario. Data for individual establishments in the ten towns and villages of Fergus, Elora, Guelph, Waterloo, Berlin, Preston, Galt, Hespeler, Paris and Brantford were coded first in mid-1982. Basic systems of industrial classification and of data entry were developed at this time. Some of the material was worked into a case-study for the Historical Atlas of Canada project, Volume II. Subsequent phases are summarized in Figure 5. As the Canadian Industry in 1871 project has evolved over several years, several combinations of coding and data entry procedures

³⁰ The distinction between primary (processing) and secondary (fabricating) manufacturing is explained in J.H. Dales, "Estimates of Canadian Manufacturing Output by Markets, 1870-1915", *Papers*, Canadian Political Science Association, Conference on Statistics 1962 and 1963, and Gilmour, *Spatial Evolution of Manufacturing*, Appendix.

Figure 5: CHRONOLOGY OF CREATING THE CANIND7I DATABASE

	I: PILOT PROJECT - TEN GRAND VALLEY TOWNS
1982	* coding procedures, including geographical and industrial classification systems * data entry, aggregation, analysis, graphics
1983	--- Accessing data for firms in 20 largest Ontario urban centres, following procedures developed in Phase 1;
1984	planning for extension to rest of Canada; applications and correspondence---
1985	II: CODING AND DATA ENTRY FOR ALL URBAN AND RURAL POWERED FIRMS IN ONTARIO (10,500) * development of dbase and mainframe procedures for editing, integration and preliminary analysis * map bases; comparison with published census data and directories; writing of papers and reports
1986	---SEPARATE PROJECT at St Mary's University sponsors entry of data for 8800 Maritime firms on Lotus spreadsheets - to Guelph by K.E. Inwood of Economics---
1987	III: EXPANSION AND CONSOLIDATION * Maritime data intensively edited to be consistent with Ontario * Rural Ontario data coded/entered * Data for all firms to include commodity details and comments * Planning for integration of whole database and for extension of project to Quebec
1988	IV: COMPLETION OF DATA ENTRY, EDITING, INTEGRATION * Development and testing of procedures to integrate all records in SAS-mainframe datasets * Intensive editing of 17000 Ontario records on mainframe * Data entry for 6000 Quebec (mainly French-language) records in Lotus * Development of project documentation: explanation of procedures; French-English dictionary of industrial vocabulary; glossary of technical terms
1989-1990	V: COMPLETION OF DATABASE: final entry and editing of Quebec; final editing and integration of Ontario and Maritime data; final versions of project documentation; refinement of SAS-mainframe procedures for handling and accessing whole database; planning for dissemination of database to other users.

have been used. Important factors in the changing project methodology have been available computer technologies and limitations of financial resources.

* The first 7,500 Ontario records were transcribed in library settings, project assistants competing with other users for the microfilm readers, and entering the details for each establishment on two 80-column coding forms. SIC codes and locational details were added in an editing stage before the data were keypunched by experienced clerks to mainframe storage.

* A further 12,500 Ontario records were coded on paper forms for keying in dBASE III to a microcomputer.

* In a separate project, the records for New Brunswick and Nova Scotia were entered directly from microfilm readers in the provincial archives to portable microcomputers. When these machine-readable records were brought to Guelph, they required very substantial and intensive editing and the addition of SIC and locational references.

* Photocopies of census schedules from microfilms borrowed on short--term loan from the National Archives had to be used for a pilot phase of Montreal records, the data being directly transcribed into the microcomputer record. (This was necessary for any work on the Quebec data until the project bought a complete set of Quebec census microfilms for 1871, which has now been given to the University of Guelph Library).

* There was even an attempt to use computer records created by the Canadian Social History Project at York University for Hamilton's industrial establishments, though this effort was frustrated by the coding conventions and technology in that project which were largely numeric. We had to add details of proprietor's names and types of industry as well as SIC and locational references and to spend a good deal of time verifying and correcting the data.

* From January 1988 to June 1989, it became possible to have the equivalent of two full-time and experienced entry and editing staff working for the project. Two work stations were set up, each combining a microfilm reader with a large enough screen for an entire census schedule to be read and a computer terminal with a monitor of high quality resolution. In this phase, about nine-tenths of the 14,500 records for the province of Quebec were entered directly from the microfilms to Lotus spreadsheets, which had the advantage over DBASEIII of resembling the census schedules in their horizontal layout. One work station was used primarily for several rounds of editing, in SAS software's "Full Screen Edit" mode on the mainframe computer, of the records that had been previously entered in the variety of software environments and methods outlined above. The mainframe computer was used for secure storage of all the datasets, for integrating the sections that had been initially entered in various formats and software programs, and also for various kinds of global editing using SAS software.

* In the second half of 1989, the remainder of the Quebec records were entered directly from the microfilms to SAS datasets on the mainframe computer. Intensive procedures were developed for verifying the basic data

and for editing the database so that it would be suitable for statistical analysis, for indexing, search and retrieval, for online transmission or copying on diskette to other users and for printing in hard copy. Particular attention was paid to the SIC designations and to verifying the numeric data for capital, employment, wages, raw materials and value of production. The mainframe computer and SAS routines that were used for very intensive tests of the consistency and accuracy of the data led to selective checks of the microfilms for five per cent of all the records. SAS routines were also used for rigorous checking and editing of the locational references and Standard Industrial Classification codes, resulting in changes to between 30 and 40 per cent of the SIC codes.

Projects of this nature are seldom well provided with resources. Archives and libraries holding the original or microfilmed records may restrict the conditions and certainly constrain the environments in which they may be consulted. Budgetary restraints of institutions and projects will reduce the availability, let alone the quality, of appropriate equipment and furniture. More importantly, they will also seriously limit the potential resources for hiring research assistants with the qualities needed to decipher, code and enter manuscript data directly into the computer record.

From our experience, we conclude that only very expert and seasoned assistants can efficiently read the microfilmed manuscript schedules and enter the data into the computer record in one operation. Such experienced staff may fully read and enter the manuscript data, assign industrial and locational codes, verify and edit at an average rate of 10 records per hour. It is a false economy to use raw recruits such as inexperienced students for such this work. In most such cases, it will be necessary for experienced editors to spend at least as long again verifying the work, comparing it with the microfilmed manuscripts and making all the corrections and additions, as it would have taken the experienced staff to enter the data themselves. Careless or ignorant errors in data entry tend to be remarkably durable and to resist several rounds of editing.³¹

If a project must depend on casual assistants, such as students working part-time during the academic year or on summer job creation schemes, it would be better to plan for three stages in the process of coding the manuscript data and making them machine-readable. Casual research assistants or even volunteers, after some careful training and following basic manuals of procedures, would transcribe data from the microfilmed census schedules to paper forms, adding codes for location and industry classification. Experienced

³¹ For those who may be considering the feasibility of large-scale projects of making routinely-generated historical records machine-readable, our labour costs may be of interest. We reckon average costs on the basis of a record having 22 standard pieces of information; there are also up to 36 additional pieces of information relating to quantities of raw materials and products for somewhat over half of all records. Including coding, data entry, editing and correction tasks, the cost per record would average about \$3.00. This does not include costs of equipment and materials, or the costs of final integration of the database in a fully consistent format.

editors would then scan these coding forms for consistency and spot check them for accuracy before approving them for data entry, which might be undertaken by the student assistants or by special data entry clerks. This procedure may be better suited to the limited availability of equipment in some projects, and has the advantage that a paper record is also retained for verification purposes.³²

In any case, there should be a thorough pilot phase at the beginning of any large-scale project, to test the efficiency and cost-effectiveness of alternative procedures. Draft manuals of procedures and classifications should be tested and finalized in this pilot phase. Assistants or volunteers should receive thorough orientation in the purposes of the project and specific training in all procedures, and an experienced editor should be available to consult with them and check their work on a daily basis for at least the first month.

Accessibility in various hardware and software environments

The Canadian Industry in 1871 has used a considerable variety of hardware and software and the CANIND71 database is not specific to any particular hardware or software. We have used mainframe and different varieties of microcomputers, and have worked in dBASEIII and Lotus for data entry and editing. We have used SAS on the University's mainframe computer for editing, calculations of derived variables and preliminary analysis. After final editing and integration of the data by the beginning of 1990, the master database is being held as SAS datasets in the CMS environment.

One of the project's objectives is to make the CANIND71 database available to other users in a variety of software environments. We expect that it will be entirely possible to distribute the database in either or both of the following formats: (1) flat files in ASCII or EBCDIC format to universities or similar bodies, with instructions as to how the whole database can be constituted for statistical analysis on the mainframe computer, and (2) high-density diskettes in DBASE for use on the microcomputer. In either format, the database will be accompanied by a manual explaining how the database was created and the meaning of the variables. We are also exploring the feasibility of transforming the database into compact disk format for CD-ROM readers.

3 DATA STRUCTURE

In this section the variables in the CANIND71 database are explained, and some of the ways in which records can be sorted and tabulated are illustrated. The database comprises variables of two main types: the **basic** ones obtained by direct transcription of information from the microfilmed manuscript schedules,

³² Darroch and Ornstein also report a preference for separating the coding and transcription stages from editing and data entry, and comment on the nature and sources of error in projects of large-scale projects handling census data. A.G. Darroch and M.D. Ornstein, "Error in Historical Data Files: A Research Note on the Automatic Detection of Error and on the Nature and Sources of Error in Coding", *Historical Methods* 12, 4 (1979): 157-168.

and the **derived** variables for which calculations or inferences were made from the first group of variables.

Basic variables in the first group are explained in relation to the format of the original schedules (see Figures 1 and 2). Variable names are printed in bold type in the following list. An alphabetically ordered explanation of variable names is also presented in Appendix A-3.

1. Kind of Industrial Establishment, Name of Proprietor or Company, and other such information / *Genres d'établissements industriels, noms de propriétaires ou des compagnies et autres renseignements*. This information is contained in the **PROPRIOR** and **TYPEEST** variables; in a small proportion of records, lengthy details that could not be fitted in the available space were placed in the **COMMENTS** field.

2. Fixed Capital invested in \$ / *Capital fixe, en piastres* - **FIXCAP**

3. Floating Capital employed in \$ / *Capital flottant, en piastres* - **FLOCAP**

4. Number of working months in the year / *Nombre de mois de travail dans l'année* - **MONTH**

5 - 8. Average Number of People employed / *Moyenne de nombre de personnes employées*:

5. Male over 16 years / *Hommes au-dessus de 16 ans* - **EMPMEN**

6. Female over 16 ans / *Femmes au-dessus de 16 ans* - **EMPWOM**

7. Boys under 16 years / *Garçons au-dessous de 16 ans* - **EMPBOY**

8. Girls under 16 years / *Filles au-dessous de 16 ans* - **EMPGIRL**

9. Aggregate amount of Yearly Wages in \$ / *Montant collectif de gages durant l'année en piastres* - **WAGES**

10. Moving Power: Kind / *Force motrice: Nature* - **TYPEPOW**

11. Moving Power: Nominal force / *Force motrice: Force nominale* (stated in units of horse power) - **FORCE**

12 - 14. Raw Materials / *Matières brutes* -- with provision for up to 12 different raw materials:

12. Kind / *Espèces* - **RAWMAT1....RAWMAT12**

Units of measurement - **RUNIT1...RUNIT12**

13. Quantities / *Quantités* - **RQUANT1...RQUANT12**

14. Aggregate Value in \$ / *Montant de la valeur en piastres* - **RVALUE1.....RVALUE12**

15 - 17. Products / *Produits* -- with provision for up to 12 different products

15 Kind / *Espèces* - **PROD1...PROD12**

Units of measurement - **PUNIT1...PUNIT12**

16. Quantities / *Quantités* - **PQUANT1...QUANT12**

17. Aggregate Value in \$ / *Montant de la valeur en piastres* -
PVALUE1.....PVALU12

18. Remarks / *Remarques* (space sometimes used for additional information, perhaps an address, or an explanation of special difficulties encountered by the enumerator; additional information also placed here occasionally by project staff) - **COMMENTS**

Four geographical variables were specified for each record:

CDID for the Census District code, in which the first digit/s were letters for the province, followed by the sequential number from 001 for Essex in southwestern Ontario to 206 for Richmond in northeastern Nova Scotia.

CDISTRIC for the name of the census district, in the coding of which qualifiers such as NORTH, SOUTH, EAST and WEST were placed after the main name.

CSD for the name of the census sub-district which often corresponds to basic municipal units of the day such as townships, villages and towns or wards of major cities. (In coding these names, qualifiers such as NORTH, etc were placed after the main name and abbreviated to N, etc. Urban municipal status was also indicated by the addition of T for Town/Ville and V for Village where applicable. In Ontario, township names which duplicated urban municipal urban names had TP added as well.

CED for the census enumerators's division, a letter and number combination for all or (more often) part of a CSD.

Derived variables were calculated as follows:

TOTEMP for the sum of all employees - men, women, boys and girls.

AVWAGE for the average monthly wage per employee, calculated only when data for any of these variables were not missing.

SUMRAWC for the sum of all RVALUs for a given establishment.

SUMPROC for the sum of all PVALUs for a given establishment.

VADD for the value added in manufacturing, the difference between SUMPROC and SUMRAWC, and calculated only when neither value was missing.

SIC for Standard Industrial Classification code, with up to 11 spaces to allow for variant suffixes and composites of two codes.

SEC for the major industry group or sector in which related basic SIC codes were grouped together, as outlined in Table 1.

PROP, in which field an entry of "F" denotes a female proprietor.

Tables on the following pages use these variable names and contain examples of some of the ways in which the basic establishment data in the CANIND71 database can be sorted, summarized and formatted. These tabulations also illustrate the generalizations expressed earlier in Figures 3 and 4 about the usefulness of coding each record by geographical location and industry class.

Figure 6 presents a sample record for an individual establishment. John Watson's enterprise making agricultural machinery in Ayr, Ontario³³ provides a good example of a record structure. It is somewhat atypical of most records, however, in including so much detail of raw materials and products. The tables that follow have been selected to show John Watson's business in its context of its place and type of industry.

Tables 2 to 5 list all the industrial establishments in the same census enumerator's division (CED), the district centred on the unincorporated village of Ayr, in North Dumfries Township, in the census district of Waterloo South. Table 2 lists all the businesses with a summary of the main data in alphabetical order of proprietor's name, Table 3 by number employed (TOTEMP), Table 4 by value of output (SUMPROC), and Table 5 by standard industrial classification (SIC) code. John Watson's business (or any other individual firm) may thus be seen in relation to others in its own community.

Tables 6 to 8 illustrate how individual records may be aggregated into successively larger geographical units. Ayr, in which John Watson had his business, was located in the second CED of North Dumfries Township. Table 6 presents the sums for most numeric variables in each of the three CEDs of North Dumfries Township. North Dumfries Township in its turn was one of seven census sub-districts (CSDs) in the Waterloo South census district; all these are listed with their summary data in Table 7. The summary industrial statistics for Waterloo South itself, in relation to the neighbouring census districts of Perth and Wellington Counties, are presented in Table 8.

The Standard Industrial Classification code of each record may also be used to group businesses of the same or related types. The more detailed SIC codes that describe several hundred industrial types may be simplified into the broader industry groups that are code-named SECs in this project. Localities, districts and regions may vary in the combinations and relative significance of industrial types. Table 9 presents the summary statistics for firms in every SIC type represented in North Dumfries Township according to the SIC codes (see Appendix A-2 for explanation of codes). Table 10 shows the statistics for

³³ John Watson's industrial enterprise has been described in John P. Shewchuk, "John Watson of Ayr", Waterloo Historical Society 74 (1986): 145-157.

Figure 6: CANIND71 SAMPLE RECORD

PROPRIOR: Watson John		TYPEEST: Foundry/Agrc Impl	
CDID: 0031	CED: C-2	CDISTRIC: Waterloo South	CSD: Dumfries N
SIC: 311	SEC: 5.14	TYPE:	MONTH: 12
FIXCAP: 15000	FLOCAP: 45000	TYPEPOW: Water	FORCE: 30
EMPMEN: 55	EMPWOM:	EMPBOY:	EMPGIRL:
TOTEMP: 55	WAGES: 17200	AVWAGE: 26.06	
SUMRAWC: 16000	SUMPROC: 55870	VADD: 39870	
RAWMAT1: Iron, Pig	RUNIT1: Ton	RQUANT1: 200	RVALUE1: 5000
RAWMAT2: Iron, Bar	RUNIT2: Ton	RQUANT1: 1000	RVALUE2: 4500
RAWMAT3: Coal	RUNIT3: Ton	RQUANT3: 100	RVALUE3: 1000
RAWMAT4: Lumber	RUNIT4:	RQUANT4:	RVALUE4: 1500
RAWMAT5: Hardware	RUNIT5:	RQUANT5:	RVALUE5: 2000
RAWMAT6: Miscellaneous Materials	RUNIT6:	RQUANT6:	RVALUE7: 2000
RAWMAT7:	RUNIT7:	RQUANT7:	RVALUE7:
RAWMAT8:	RUNIT8:	RQUANT8:	RVALUE8:
RAWMAT9:	RUNIT9:	RQUANT9:	RVALUE9:
RAWMAT10:	RUNIT10:	RQUANT10:	RVALUE10:
RAWMAT11:	RUNIT11:	RQUANT11:	RVALUE11:
RAWMAT12:	RUNIT12:	RQUANT12:	RVALUE12:
PROD1: Threshing Machines	PUNIT1:	PQUANT1: 23	PVALUE1: 920
PROD2: Reaping, Mowing Machines	PUNIT2:	PQUANT2: 200	PVALUE2: 28000
PROD3: Ploughs	PUNIT3:	PQUANT3: 500	PVALUE3: 7500
PROD4: Grain Drills	PUNIT4:	PQUANT4: 30	PVALUE4: 2100
PROD5: Hay Rakes	PUNIT5:	PQUANT5: 40	PVALUE5: 1400
PROD6: Cultivators, Two Horse	PUNIT6:	PQUANT6: 30	PVALUE7: 750
PROD7: Straw Cutters	PUNIT7:	PQUANT7: 50	PVALUE7: 2000
PROD8: Root Cutters	PUNIT8:	RQUANT8: 50	PVALUE8: 1500
PROD9: Cultivators, One Horse	PUNIT9:	PQUANT9: 50	PVALUE9: 500
PROD10: Seed Drills	PUNIT10:	PQUANT10: 30	PVALUE10: 600
PROD11: Field Rollers	PUNIT11:	PQUANT11: 10	PVALUE11: 320
PROD12: Miscellaneous Castings	PUNIT12:	PQUANT12:	PVALUE12: 2000
COMMENTS:			

Table 2

INDUSTRIAL ESTABLISHMENTS IN AYR DISTRICT BY PROPRIETOR NAME

CDD	CED	PROPRIETOR	TYPEEST	SIC	FIXCAP	TYPEPOW	FORCE	TOTEMP	WAGES	SUMRAWC	SUMPROC	VADD
0031	C 2	BAKER WILLIAM	BOOT/SHOE SHOP	174	300	.	.	2	700	800	1728	928
0031	C 2	BELL & DRYDEN	WAGON SHOP	329	.	.	.	4	1250	2050	4800	2750
0031	C 2	BUCKLEY TIMOTHY	TIN SHOP	304-T	.	.	.	1	100	300	680	380
0031	C 2	CAMPBELL ROBERT	SHOE SHOP	174	150	.	.	1	300	200	900	700
0031	C 2	CUNNINGHAM JANE	BREWERY	109-B	3000	HORSE	2	2	348	600	1200	600
0031	C 2	CUTHBERTSON HUGH	BOOT/SHOE SHOP	174	.	.	.	4	700	1500	2550	1050
0031	C 2	GILLESPIE JOSEPH	PAINT SHOP	421-D	250	.	.	1	350	440	1200	760
0031	C 2	GOLDIE DAVID	GREENFIELD MILLS	105	20000	WATER	80	7	3000	125000	131000	6000
0031	C 2	GOLDIE DAVID	STAVE FCY/COOPER	251/259-C	1500	WATER	10	11	2750	2587	9150	6563
0031	C 2	HENDERSON R & A	BLACKSMITH SHOP	896	350	.	.	1	240	650	1800	1150
0031	C 2	HOPE THOMAS	CABINET SHOP/COOP	261/258	400	HORSE	1	1	336	400	1036	636
0031	C 2	KAY JAMES	CARPENTER SHOP	259	2000	WATER	6	3	800	800	2000	1200
0031	C 2	MATHIESON ALEX	WAGON SHOP	329	650	.	.	1	150	79	440	361
0031	C 2	MCKNIGHT THOMAS	HAIRNESS SHOP	179-S	.	.	.	1	325	500	1200	700
0031	C 2	MURRAY JOHN	BAKERY	107/108-C	3000	.	.	2	400	1500	2175	675
0031	C 2	PIPER JAMES	SAW MILL	251	3000	WATER	20	2	300	1000	2000	1000
0031	C 2	PIPER JAMES	WOOLEN FACTORY	182	12000	WATER	20	18	3600	9000	15500	6500
0031	C 2	PIPER JAMES	FLOUR/FEED MILL	105	12000	WATER	20	8	2000	60000	68000	8000
0031	C 2	PIPER JAMES	OATMEAL MILL	105-O	6000	WATER	6	2	600	6000	8000	2000
0031	C 2	PIPER JAMES	COOPER SHOP	259-C	500	.	.	4	1600	2400	4700	2300
0031	C 2	SHEPHERD ALEX	HAIRNESS SHOP	179-S	45	.	.	1	300	660	1250	590
0031	C 2	WATSON JOHN	FOUNDRY/AGRC IMPL	311	15000	WATER	30	55	17200	16000	55870	39870
0031	C 2	WILTON ROBERT	BLACKSMITH SHOP	896	400	.	.	1	75	300	900	600

Table 3

INDUSTRIAL ESTABLISHMENTS IN AYR DISTRICT BY NO. EMPLOYED

CDD	CED	PROPRIETOR	TYPEEST	SIC	FIXCAP	TYPEPOW	FORCE	TOTEMP	WAGES	SUMRAWC	SUMPROC	VADD
0031	C 2	WATSON JOHN	FOUNDRY/AGRC IMPL	311	15000	WATER	30	55	17200	16000	55870	39870
0031	C 2	PIPER JAMES	WOOLEN FACTORY	182	12000	WATER	20	18	3600	9000	15500	6500
0031	C 2	GOLDIE DAVID	STAVE FCY/COOPER	251/259-C	1500	WATER	10	11	2750	2587	9150	6563
0031	C 2	PIPER JAMES	FLOUR/FEED MILL	105	12000	WATER	20	8	2000	60000	68000	8000
0031	C 2	GOLDIE DAVID	GREENFIELD MILLS	105	20000	WATER	80	7	3000	125000	131000	6000
0031	C 2	CUTHBERTSON HUGH	BOOT/SHOE SHOP	174	.	.	.	4	700	1500	2550	1050
0031	C 2	BELL & DRYDEN	WAGON SHOP	329	.	.	.	4	1250	2050	4800	2750
0031	C 2	PIPER JAMES	COOPER SHOP	259-C	500	.	.	4	1600	2400	4700	2300
0031	C 2	MATHIESON ALEX	CARPENTER SHOP	259	2000	WATER	6	3	800	800	2000	1200
0031	C 2	PIPER JAMES	SAW MILL	251	3000	WATER	20	2	300	1000	2000	1000
0031	C 2	CUNNINGHAM JANE	BREWERY	109-B	3000	HORSE	2	2	348	600	1200	600
0031	C 2	PIPER JAMES	OATMEAL MILL	105-O	6000	WATER	6	2	600	6000	8000	2000
0031	C 2	MURRAY JOHN	BAKERY	107/108-C	3000	.	.	2	400	1500	2175	675
0031	C 2	BAKER WILLIAM	BOOT/SHOE SHOP	174	300	.	.	2	700	800	1728	928
0031	C 2	KAY JAMES	CABINET SHOP/COOP	261/258	400	HORSE	1	1	200	75	660	585
0031	C 2	MCKNIGHT THOMAS	WAGON SHOP	329	650	.	.	1	150	79	440	361
0031	C 2	SHEPHERD ALEX	HAIRNESS SHOP	179-S	45	.	.	1	300	660	1250	590
0031	C 2	MURRAY JOHN	HAIRNESS SHOP	179-S	.	.	.	1	325	500	1200	700
0031	C 2	WILTON ROBERT	BLACKSMITH SHOP	896	400	.	.	1	75	300	900	600
0031	C 2	GILLESPIE JOSEPH	PAINT SHOP	421-D	250	.	.	1	350	440	1200	760
0031	C 2	CAMPBELL ROBERT	SHOE SHOP	174	150	.	.	1	300	200	900	700
0031	C 2	BUCKLEY TIMOTHY	TIN SHOP	304-T	.	.	.	1	100	300	680	380
0031	C 2	HOPE THOMAS	BLACKSMITH SHOP	896	350	.	.	1	336	400	1036	636
0031	C 2	HENDERSON R & A	BLACKSMITH SHOP	896	.	.	.	1	240	650	1800	1150
					80545						318739	85898
					195						37624	232841
					134						6000	6000

Table 4

INDUSTRIAL ESTABLISHMENTS IN AYR DISTRICT BY OUTPUT

CDD	CED	PROPRIOR	TYPEEST	SIC	FIXCAP	TYPEPOW	FORCE	TOTEMP	WAGES	SUMRAWC	SUMPROC	VADD
0031	C-2	GOLDIE DAVID	GREENFIELD MILLS	105	20000	WATER	80	7	3000	125000	131000	6000
0031	C-2	PIPER JAMES	FLOUR/FEED MILL	105	12000	WATER	20	8	2000	60000	68000	8000
0031	C-2	WATSON JOHN	FOUNDRY/AGRC IMPL	311	15000	WATER	30	55	17200	16000	55870	39870
0031	C-2	PIPER JAMES	WOOLEN FACTORY	182	12000	WATER	20	18	3600	9000	15500	6500
0031	C-2	GOLDIE DAVID	STAVE FCY/COOPER	251/259-C	1500	WATER	10	11	2750	2587	9150	6563
0031	C-2	PIPER JAMES	OATMEAL MILL	105-O	6000	WATER	6	2	600	6000	8000	2000
0031	C-2	BELL & DRYDEN	WAGON SHOP	329	500	.	.	4	1250	2050	4800	2750
0031	C-2	PIPER JAMES	COOPER SHOP	259-C	.	.	.	4	1600	2400	4700	2300
0031	C-2	CUTHBERTSON HUGH	BOOT/SHOE SHOP	174	3000	.	.	4	700	1500	2550	1050
0031	C-2	MURRAY JOHN	BAKERY	107/108-C	3000	WATER	20	2	300	1000	2000	1000
0031	C-2	PIPER JAMES	SAW MILL	251	2000	WATER	6	3	800	800	2000	1200
0031	C-2	MATHIESON ALEX	CARPENTER SHOP	259	300	.	.	1	240	650	1800	1150
0031	C-2	HENDERSON R & A	BLACKSMITH SHOP	896	45	.	.	2	700	800	1728	928
0031	C-2	BAKER WILLIAM	BOOT/SHOE SHOP	174	3000	HORSE	2	1	300	660	1250	590
0031	C-2	SHEPHERD ALEX	HARNESS SHOP	179-S	250	.	.	1	348	600	1200	600
0031	C-2	CUNNINGHAM JANE	BREWERY	109-B	350	.	.	1	350	500	1200	700
0031	C-2	MORTON ALEX	HARNESS SHOP	179-S	150	.	.	1	325	440	1200	760
0031	C-2	GILLESPIE JOSEPH	PAINT SHOP	421-D	300	.	.	1	336	400	1036	636
0031	C-2	HOPE THOMAS	BLACKSMITH SHOP	896	400	.	.	1	75	300	900	600
0031	C-2	WHITSON ROBERT	SHOE SHOP	174	150	.	.	1	300	200	900	700
0031	C-2	CAMPBELL ROBERT	TIN SHOP	304-T	400	HORSE	1	1	100	300	680	380
0031	C-2	BUCKLEY TIMOTHY	CABINET SHOP/COFF	261/258	650	.	.	1	200	75	660	585
0031	C-2	KAY JAMES	WAGON SHOP	329	.	.	.	1	150	79	440	361

Table 5

INDUSTRIAL ESTABLISHMENTS IN AYR DISTRICT BY SIC CODE

CDD	CED	PROPRIOR	TYPEEST	SIC	FIXCAP	TYPEPOW	FORCE	TOTEMP	WAGES	SUMRAWC	SUMPROC	VADD
0031	C-2	GOLDIE DAVID	GREENFIELD MILLS	105	20000	WATER	80	7	3000	125000	131000	6000
0031	C-2	PIPER JAMES	FLOUR/FEED MILL	105	12000	WATER	20	8	2000	60000	68000	8000
0031	C-2	PIPER JAMES	OATMEAL MILL	105-O	6000	WATER	6	2	600	6000	8000	2000
0031	C-2	MURRAY JOHN	BAKERY	107/108-C	3000	.	.	2	400	1500	2175	675
0031	C-2	CUNNINGHAM JANE	BREWERY	109-B	3000	HORSE	2	2	348	600	1200	600
0031	C-2	CUTHBERTSON HUGH	BOOT/SHOE SHOP	174	300	.	.	4	700	1500	2550	1050
0031	C-2	BAKER WILLIAM	BOOT/SHOE SHOP	174	45	.	.	2	700	800	1728	928
0031	C-2	CAMPBELL ROBERT	SHOE SHOP	174	150	.	.	1	300	200	900	700
0031	C-2	SHEPHERD ALEX	HARNESS SHOP	179-S	300	.	.	1	300	660	1250	590
0031	C-2	MORTON ALEX	HARNESS SHOP	179-S	45	.	.	1	325	500	1200	700
0031	C-2	PIPER JAMES	WOOLEN FACTORY	182	12000	WATER	20	18	3600	9000	15500	6500
0031	C-2	PIPER JAMES	SAW MILL	251	3000	WATER	20	2	300	1000	2000	1000
0031	C-2	GOLDIE DAVID	STAVE FCY/COOPER	251/259-C	1500	WATER	10	11	2750	2587	9150	6563
0031	C-2	MATHIESON ALEX	CARPENTER SHOP	259	2000	WATER	6	3	800	800	2000	1200
0031	C-2	PIPER JAMES	COOPER SHOP	259-C	500	.	.	4	1600	2400	4700	2300
0031	C-2	KAY JAMES	CABINET SHOP/COFF	261/258	400	HORSE	1	1	200	75	660	585
0031	C-2	BUCKLEY TIMOTHY	TIN SHOP	304-T	15000	WATER	30	55	17200	16000	55870	39870
0031	C-2	WATSON JOHN	FOUNDRY/AGRC IMPL	311	.	.	.	4	1250	2050	4800	2750
0031	C-2	BELL & DRYDEN	WAGON SHOP	329	650	.	.	1	150	79	440	361
0031	C-2	MCKNIGHT THOMAS	WAGON SHOP	329	250	.	.	1	350	440	1200	760
0031	C-2	GILLESPIE JOSEPH	PAINT SHOP	421-D	.	.	.	1	240	650	1800	1150
0031	C-2	HENDERSON R & A	BLACKSMITH SHOP	896	350	.	.	1	336	400	1036	636
0031	C-2	HOPE THOMAS	BLACKSMITH SHOP	896	400	.	.	1	75	300	900	600
0031	C-2	WHITSON ROBERT	BLACKSMITH SHOP	896	.	.	.	1	150	79	440	361
										=====	=====	=====
					80545				37624	232841	318739	85898
								195	134	=====	=====	=====

Table 9

INDUSTRIAL STRUCTURE OF NORTH DUMFRIES TOWNSHIP, 1871

SIC	OBSERV	FORCE	FIXCAP	FLOCAP	EMPMEN	EMPWOM	EMPBOY	EMPGIRL	TOTEMP	WAGES	SUMRAWC	SUMPROC	VADD
104	3	.	3770	4860	3	5	4	1	13	800	5812	7020	1208
105	3	112	36000	90000	17	.	.	.	17	5600	190000	206200	16200
105-O	1	6	6000	5000	2	.	.	.	2	600	6000	8000	2000
107	1	.	3000	500	2	.	.	.	2	400	1500	2175	675
109-B	1	2	3000	1000	2	.	.	.	2	348	600	1200	600
109-C	1	6	300	.	3	.	.	.	3	180	1575	2800	1225
174	4	.	460	2610	8	.	.	.	8	1740	2780	5738	2958
179-S	2	.	45	1475	2	.	.	.	2	625	1160	2450	1290
182	1	20	12000	10000	10	6	2	.	18	3600	9000	15500	6500
182-W	3	.	495	.	4	3	2	1	10	200	674	1145	471
239	1	.	200	150	2	2	.	.	4	300	200	800	600
251	5	85	8900	15750	33	.	9	.	42	6010	16337	32745	16408
251-S	1	8	300	10	1	.	.	.	1	50	41	120	79
259	2	6	2200	1630	4	.	.	.	4	880	816	2100	1284
259-C	1	.	500	1500	4	.	.	.	4	1600	2400	4700	2300
261	1	1	400	250	1	.	.	.	1	200	75	660	585
304-T	1	.	600	600	1	.	.	.	1	100	300	680	380
311	1	30	15000	45000	55	.	.	.	55	17200	16000	55870	39870
315	1	6	1000	500	4	.	.	.	4	1800	500	2500	2000
315-P	2	8	380	30	2	.	.	.	2	250	61	560	499
329	4	.	1230	3500	8	.	.	.	8	1650	2285	6340	4055
329-P	1	.	300	1000	3	.	.	.	3	500	600	1500	900
421-D	1	.	250	700	1	.	.	.	1	350	440	1200	760
896	8	.	2450	2556	10	.	.	.	10	1911	2225	6786	4561
	=====	=====	=====	=====	=====	=====	=====	=====	=====	=====	=====	=====	=====
	50	290	98180	188621	182	16	17	2	217	46894	261381	368789	107408

Table 10

INDUSTRIAL STRUCTURE OF NORTH DUMFRIES TOWNSHIP

12:01 SATURDAY, JANUARY 13, 1990

CSD=DUMFRIES N

		OBSERV		FORCE		FIXCAP		TOTEMP		SUMPROC		VADD	
		SUM	PERC- ENT	SUM	PERC- ENT	SUM	PERC- ENT	SUM	PERC- ENT	SUM	PERC- ENT	SUM	PERC- ENT
CSD	SEC												
DUMFRIES N	5.01	10	20.0	126	43.4	52070	53.0	39	18.0	227395	61.7	21908	20.4
	5.04	6	12.0	.	.	505	0.5	10	4.6	8188	2.2	4248	4.0
	5.05	4	8.0	20	6.9	12495	12.7	28	12.9	16645	4.5	6971	6.5
	5.06	1	2.0	.	.	200	0.2	4	1.8	800	0.2	600	0.6
	5.08	9	18.0	99	34.1	11900	12.1	51	23.5	39665	10.8	20071	18.7
	5.09	1	2.0	1	0.3	400	0.4	1	0.5	660	0.2	585	0.5
	5.13	1	2.0	1	0.5	680	0.2	380	0.4
	5.14	4	8.0	44	15.2	16380	16.7	61	28.1	58930	16.0	42369	39.4
	5.15	5	10.0	.	.	1530	1.6	11	5.1	7840	2.1	4955	4.6
	6	1	2.0	.	.	250	0.3	1	0.5	1200	0.3	760	0.7
	10	8	16.0	.	.	2450	2.5	10	4.6	6786	1.8	4561	4.2
ALL		50	100.0	290	100.0	98180	100.0	217	100.0	368789	100.0	107408	100.0

Table 11
INDUSTRY CLASSES IN THE MACHINERY SECTOR, 1871

SIC	OBSERV	FORCE	FIXCAP	FLOCAP	EMP MEN	EMP WOM	EMP BOY	EMPGIRL	TOTEMP	WAGES	SUMRAWC	SUMPROC	VADD
311	257	1562	717550	864858	1703	3	75	.	1781	573245	520452	1778688	1236536
311/105	1	80	42000	40000	90	.	.	.	90	35000	43000	119000	76000
311/254	6	57	17800	13251	33	.	4	.	37	11900	7426	29600	22174
311/259	2	4	2040	100	4	.	.	.	4	1600	100	3000	2900
311/261	4	20	2340	2270	5	.	.	.	5	1050	1255	3350	2095
311/294	9	63	25400	8300	36	.	.	.	36	10660	12234	27110	14876
311/301	1	40	40000	18000	25	.	.	.	25	6000	14650	31000	16350
311/304-T	1	.	5000	1500	4	.	.	.	4	1200	800	2300	1500
311/306	1	16	2000	1500	2	.	1	.	3	800	110	1600	1490
311/307-S	13	153	126500	149200	236	.	7	.	243	83250	96436	235004	136568
311/309	3	2	400	1750	7	.	.	.	7	1850	1043	3105	2062
311/315	18	355	307800	264750	519	10	27	.	556	197056	247725	659590	407565
311/315-E	4	104	62000	124600	161	.	7	.	168	63800	61200	203550	142350
311/315-P	1	4	1000	300	2	.	1	.	3	450	225	850	625
311/329	24	39	21945	12830	61	.	2	.	63	13158	10630	42982	32352
311/329-P	1	.	100	500	2	.	.	.	2	450	250	1000	750
311/421-C	2	.	350	120	2	.	.	.	2	500	140	900	760
311/896	16	.	4400	7775	31	.	1	.	32	8930	6587	22101	15514
315	38	363	280350	1165600	285	2	7	.	294	102846	162579	367600	206321
315-E	4	105	68000	78500	110	.	11	.	121	40300	27500	155000	127500
315-E/301	7	204	166000	148200	319	.	23	3	345	102847	94797	292240	197443
315-E/302	1	12	3000	800	7	.	.	.	7	2550	770	5400	4630
315-E/308	1	8	5000	1000	9	.	.	.	9	7000	2400	2500	100
315-E/311	2	60	202000	260000	264	.	16	.	280	129160	59375	311000	251625
315-E/315	6	160	189700	205000	399	.	6	.	405	128873	64500	367800	303300
315-E/326	1	50	120000	200800	230	.	1	.	230	80000	112900	275600	162700
315-H	8	10	2400	3170	10	1	1	.	12	1230	365	4932	4167
315-H/259	2	.	400	50	2	.	.	.	2	420	19	700	681
315-H/259-W	3	12	1450	750	3	.	.	.	3	1200	165	2210	2045
315-H/261	2	9	1080	200	3	.	1	.	4	1140	110	1160	1050
315-H/329-P	2	12	300	110	2	.	.	.	2	490	38	628	590
315-M	2	18	12000	12320	28	.	.	.	28	10400	7800	23000	15200
315-M/315	1	35	30000	26000	80	.	.	.	80	30000	11640	60000	48360
315-P	144	209	54464	41301	238	.	7	.	245	50843	27729	156332	127503
315-P/259	3	22	1200	400	4	.	.	.	4	750	455	2600	2145
315-P/259-C	1	.	500	200	2	.	.	.	2	700	300	1800	1500
315-P/261	4	8	2150	440	7	.	.	.	7	2045	662	3026	2364
315-P/311	3	26	6100	11200	14	2	.	.	16	5900	2550	16400	13850
315-S	11	168	215500	220400	695	16	179	.	890	314856	284825	983231	698406
315-S/315	1	4	2200	800	1	.	1	.	2	800	300	1760	1460
315/294	2	18	8500	2400	13	.	.	.	13	4000	1600	7500	5900
315/306	2	4	200	50	3	.	.	.	3	700	100	1250	250
315/307-S	1	5	10000	8000	10	.	.	.	10	5500	6300	30000	23700
315/311	12	139	55895	37602	97	.	10	.	107	34468	28100	90805	62705
315/326	1	60	100000	500000	200	.	.	.	200	100000	35000	620000	585000
315/327	1	40	10000	2000	16	.	.	.	16	7300	5800	16000	10200
318	2	20	8000	8000	25	.	4	.	29	9000	6320	30750	24430
=====										=====	=====	=====	=====
632	632	4280	2935014	4446897	5999	34	391	3	6427	2186217	1969262	6995954	4997592

Table 1.2

CANADA: MACHINERY MANUFACTURERS WITH OUTPUT OVER \$50,000

CDID	CED	PROPRIOR	TYPEEST	SIC	FIXCAP	TYPEPOW	FORCE	TOTEMP	WAGES	SUMRAWC	SUMPROC	VADD
0047	A-3	HAMILTON WILLIAM	FOUNDRY/MACHINE S	315/326	100000	STEAM	60	200	100000	35000	620000	585000
0048	E-2	HALL JOSEPH & CO	ENGINES/AGRC IMPL	315-E/311	200000	STEAM	50	266	125000	56600	300000	243400
0046	C-1	DICKEY NEIL &	ENGINES/FOUNDRY	315-E/326	120000	STEAM	50	230	80000	112900	275600	162700
0024	E-2	WANZER R M & C	SEWING MACHINE MF	315-S	58000	STEAM	40	275	86000	65000	210000	145000
0033	C-3	GUELPH SEWING M C	SEWING MACHINE FC	315-S	65000	STEAM	12	180	75000	62000	204800	142800
0024	A-1	WILSON BOWMAN & C	SEWING MACHINE MF	315-S	8000	STEAM	15	131	60000	50000	200000	150000
0031	D-3	GOLDIE/MCCULLOCH/	ENGINE/MACHINE WO	315-E/315	100000	STEAM	50	203	65000	27700	180000	152300
0029	E-3	THOMPSON & WILLIA	AGRC WORKS/ENGINE	311/315-E	25000	STEAM	30	106	40000	50000	150000	100000
Q106	A-9	BARTLEY W P & CO	ENGINE WORKS/FOUN	315-E/315	150000	WATER	160	222	49200	36250	128175	91925
0024	D-2	SAWYER S D & C	AGRC IMPLEMENTS	311	10000	STEAM	22	90	45000	73000	125000	52000
NS19	A-1	MONTGOMERY WILLIA	MACHINE SHOP/FOUN	315-E/315	55480	STEAM	30	63	21181	15540	125000	109460
0015	D-2	WATEROUS C H & C	ENGINES/MACHINE	315-E/315	57000	STEAM	40	118	40573	19700	120000	100300
Q106	A-9	GILBERT E E	ENGINE/MACHINE WO	315-E/315	35000	STEAM	40	145	50000	46500	120000	73500
0044	B-5	ABELL JOHN	AGRC WKS/FOUR MI	311/105	42000	STEAM	80	90	35000	43000	119000	76000
0079	B	FROST & WOOD	FOUNDRY/AGRC IMPL	311/315	28000	STEAM	30	74	26000	55000	115000	60000
0044	B-2	PATTERSON BROS	AGRC MACHINE WORK	311	50000	W/S	30	125	50000	30650	113630	82980
Q106	A-3	LAWLOR J D	SEWING MACHINE MF	315-S	59000	STEAM	4	51	25000	12500	112500	100000
0013	F-2	NOXON BROS	FOUNDRY	311/315	70000	STEAM	60	103	42000	50500	111400	60900
0079	B	COSSITT & BRO	FOUNDRY/AGRC IMPL	311/315	50000	WATER	60	65	25000	27700	107000	79300
Q106	A-9	MCDUGALL JOHN	FOUNDRY/MILLWRIGHT	315/315-E	30000	WATER	25	140	60000	37000	104000	67000
0039	E-2	HAGGERT BROS	FOUNDRY/AGRC IMPL	311/307-S	30000	STEAM	20	93	31500	50500	103500	53000
0013	F-2	EASTWOOD & CO	AGRC IMPLEMENT MF	311/315	30000	STEAM	16	70	28000	51860	101000	49140
0024	A-1	GARDNER & CO	SEWING MACHINE MF	315-S	6000	STEAM	20	70	8000	25000	100000	75000
0024	C-2	BECKETT F G & C	ENGINE/BOILER MF	315-E/301	100000	STEAM	50	120	40000	40000	100000	60000
0033	C-3	THURESSON EYRE ?	SEWING MACHINE FC	315-S	25000	STEAM	15	70	30000	42000	85000	43000
0022	E-3	MASSEY MF CO	CLOTH CARD MF	315	20000	WATER	10	8	3050	80500	84800	4300
0050	E	BRUSH GEORGE	AGRC IMPLEMENTS	311	20000	STEAM	30	61	20000	18845	80000	61155
Q106	A-2	LOCKMAN & WILSON/	FOUNDRY/MACHINE S	315-E	50000	STEAM	20	80	25000	25000	80000	55000
0034	D	LEVY CHARLES & C	STEAM ENGINE FACT	315-S	15000	STEAM	25	75	22000	20000	75000	55000
0046	B-1	PAXTON & TATE & C	IRON FOUNDRY	311/315	40000	STEAM	30	46	10000	12000	70000	58000
0049	B-3	FLEURY JOSEPH	AGRC IMPLEMENT MF	311	12000	STEAM	20	55	20000	14000	65000	51000
0043	C	WILLIAMS C W & CO	SEWING MACHINE MF	315-S	40000	STEAM	8	90	31200	28760	63280	34520
Q106	C-1	COPP ANTHONY & WM	IRON FOUNDRY	311/307-S	40000	STEAM	16	60	23000	10000	62400	52400
0024	B-2	MCKECHNIE & BERTR	CANADA TOOL WORKS	315-M/315	30000	STEAM	35	80	30000	24800	62000	37200
0023	C-2	HAMILTON WILLIAM	FOUNDRY/MACHINE F	315-E/301	20000	W/S	25	40	6700	12260	60000	48360
0056	C-1	GARDNER ROBERT &	MACHINE MANUFACTO	315	36000	STEAM	12	32	12000	9000	60000	47740
Q106	A-3	BEAUCHEMIN MOISE	FONDERIE	311	40300	STEAM	10	40	14500	12200	60000	51000
Q120	F-3	BANNER S M CO	SEWING MACHINE MF	315-S	5000	WATER	20	25	10000	12200	58500	46300
Q140	A-2	LAWLOR J D	SEWING MACHINE MF	315-S	83500	STEAM	10	50	24000	12000	57600	45600
Q104	A-1	ELLIOTT JOHN	AGRC IMPLEMENT MF	311	12000	STEAM	12	34	16000	6000	57500	51500
0010	C-2	BROWN & PATTERSON	FOUNDRY/MACHINE S	311	20000	STEAM	25	50	23500	6500	56000	49500
0048	C-1	WATSON JOHN	FOUNDRY/AGRC IMPL	311	15000	WATER	30	55	17200	18100	56000	37900
0031	C-2	WILSON THOMAS & C	DUNDAS FOUNDRY	315-E/301	14000	W/S	60	76	19167	16000	55870	39870
0023	C-2	GARTH CHARLES & C	DOMINION METAL WO	315-E	40000	STEAM	25	86	13000	27237	55000	27763
Q104	B	HYSLIP & RONALD	STEAM ENGINE FACT	315-E	15000	STEAM	50	42	18000	26000	55000	29000
0002	G-2	HARRIS A & SON	AGRC IMPLEMENT MF	311	5000	STEAM	15	21	8000	6000	50000	44000
0021	E-2	GATES G W & CO	SEWING MACHINE MF	315-S	10000	STEAM	20	25	12000	12200	50000	37800
0046	B-1	BESSET	FOUNDRY/ENGINEER	315	20000	STEAM	10	52	10000	10000	50000	40000
Q145	A-2				2126280		1557	4538	1643771	1538942	5504555	3965613

selected variables in the broader SEC groupings (which are defined in Table 1 and Appendix A-2).

By sorting on the SIC and SEC codes, users of the CANIND71 database may link businesses of a certain industrial type to firms of the same type in other localities and regions. Table 11 summarizes the numeric data for all Ontario businesses in the machinery sector (SEC 5.14) to which John Watson's firm belonged. The large number of composite SIC codes in this sector should be noted, as many enterprises were diversified rather than specialized but made several lines of products. In Table 12, we can see how John Watson's business ranked in relation to other Canadian firms in the machinery manufacturing sector. Of the 49 firms of this type with an output of at least \$50,000 in 1870-1, John Watson's ranked 42nd throughout the four provinces.

4 CODING AND EDITING PROCEDURES AND PROBLEMS

Clearly, meticulous accuracy and intelligent consistency were crucial in the transcription process. We tried hard to be consistent both with respect to what was actually written and how the information might be interpreted at the various levels. We cannot overstate the importance of careful, alert and intelligent transcribers in the project. We came to realize the value of considering the various details recorded on each establishment in relation to one another, and not just as columns of separate data. For example, if an assistant was having problems interpreting the data in one column, it was helpful to refer to the other related columns. Considering the aggregate value of raw materials in relation to the aggregate value of products was also useful. Linking the ratio of the number of employees to the amount of wages paid and the number of months worked alerted the transcriber to any gross errors which might have been made by the enumerator, or could be a safeguard against transcriber error. Viewing the data in their context could be valuable not only at the level of the individual establishment, but similar types of industrial establishments within a certain census enumerator's district could also be viewed en bloc. It was vital not to miss any of the microfilmed census schedules, and a valuable guide in preventing this the shelf-list of 1871 microfilms produced by the National Archives.³⁴

Common errors in transcription and data entry of handwritten manuscript material were found to occur with certain letters. In Victorian calligraphy, "S" resembled "F", "M" was hard to distinguish from "N" or "W", and a double "SS" was handwritten to resemble a "P", so that "Jane Wissler" looked like "Jane Wipler". One had to be alert for inconsistencies by the enumerators, for example in the order in which they gave surname and forename of proprietors.

³⁴ Public Archives of Canada, Finding Aid Census 1871 Recensement Instrument de Recherche (Record Group 31). We are grateful to T.W. Hillman for advice on the 1871 census microfilms. The shelf-list section of the finding aid indicates on which microfilm reel a particular census enumerator's division may be found. It also indicates in which divisions schedules are missing, incomplete or have no entry. While this was not completely accurate (in that we did find some schedules that were declared to be missing) it was a most useful reference aid.

Errors in numeric values could arise when an assistant misinterpreted the occasional use of a double zero by an enumerator to indicate cents as another pair of zeros for the number of dollars. Thus what was intended as \$500⁰⁰ might easily be transcribed and coded as \$50000. Fractions were sometimes used on the manuscript schedules in stating the number of working months in the year and quantities of raw materials and products. They were entered in the computer record as decimals, as in the following examples:

two weeks = 0.5 month
one and a half tons = 1.5 tons

At the end of transcription and data entry for a whole census district, the computer record was printed out and closely edited against the microfilm. Points queried by transcribers were looked at carefully for further verification. The SIC codes were checked against the TYPEEST and the raw materials and products columns. Generally, the whole set of data was scanned for oddities and inconsistencies, such as typographical errors, textual inconsistencies, proper agreement between adjectives and nouns, proper spacing of comments, multiple raw materials and products, and so on. Industrial data in all districts were checked at least against the microfilm. Finally, various global editing procedures were performed in SAS on the mainframe computer, to check for consistency of format in all fields, for possible anomalies that might require further microfilm checks and, especially, for the accuracy of the Standard Industrial Classification codes.

We comment further here on project practice with several variables, in which certain rules were followed in the interest of consistency and brevity.³⁵

Proprior The surname of the proprietor was entered first followed by the forename. If there was not enough room on the spreadsheet or coding form for the names of all partners or the full corporate name, additional details would be added in the COMMENTS field. Some enumerators were inconsistent in the order in which they gave surnames and forenames. Abbreviations of the following words, Junior(Jr), Senior(Sr), Brother(Bro), Brothers(Bros), were made so as not to confuse the user who might mistakenly interpret them as a surname if typed in full. If the manuscript census was particularly difficult to read or illegible, the assistant simply provided his or her best interpretation of the proprietor's name or at the very least the first initial of the surname or forename followed by a question mark (?).

In order to achieve the best possible record of industrial activity in 1871, and to overcome some problems of calligraphy in the manuscript schedules, we checked proprietors' names and types of establishment and products in some districts and urban centres. We did this for all the urban and "proto-urban"

³⁵ This section draws upon examples and points in the Procedures Manual prepared by the project's senior research assistant Janine Grant in January 1988 and revised January 1989.

places in Ontario,³⁶ for the three census districts of Montreal Centre, Montreal East and Montreal West, for selected rural areas in Ontario, and as spot checks in difficult cases throughout the four provinces. Information from the census schedules was checked against lists in the R.G. Dun reference books of credit-worthiness and in city, county and provincial directories and historical atlases. In cases where the census enumerator's spelling of the proprietor's name differed significantly from that in a contemporary printed source, while clearly referring to the same establishment, the record was "corrected" if this would result in a more plausible or conventional rendering of the name. Such comparisons against other contemporary sources also helped to identify a few omissions of industrial businesses from the census record.

Typeest The staff of census enumerators in 1871 were not provided with any controlled vocabulary to use in describing the types of industrial activity they encountered. They entered information about the kinds of industry in the language that came naturally to them or to their informants. Thus there is considerable variety in the terms used to describe the workplaces and the industrial processes carried on there.

For one example, the establishments in which bread was baked were described in the following ways in English:

Bake house, Bake shop, Baker, Baker factory, Baker house, Baker shop, Baker & confectioner, Bakery, Bakery establishment, Bakery shop, Bakery & biscuit factory, Bakery & confectionery, Bakery & crackers, Bakery & pastry, Baking establishment, Baking manufactory, Bread bakery, Bread & biscuit bakery, Bread & cake establishment, Bread & fancy bakery, Confectionery, Confectionery & bakery, and Soft bread bakery.

The activity of making shoes was recorded in the following kinds of establishment by francophone enumerators:

Atelier de chaussure, Boutique de chaussure, Fabrique de chaussure, Boutique de chaussures, Manufacture de chaussures, Cordonnerie, Atelier de cordonnerie, Boutique de cordonnerie, Maison de cordonnerie, Cordonnier, Atelier de cordonnier, Boutique de cordonnier, Boutique de cordonniere, Fabrique de souliers, Manufacture de souliers et bottes, Manufacture de souliers.

At first, we considered standardizing all such variants into one term if their product was the same. But, for several reasons, it was decided that the natural language in all its variations should be transcribed with minimal changes into the computer record. One factor was our wish to reduce the potential for error in data entry. The task of deciphering and transcribing data from the microfilmed manuscript schedules was demanding enough without asking assistants to make such judgments at the same time. Secondly, we considered that the natural language, whether in English or French, would have intrinsic interest for some users of the database. Usages of the terms to describe workplaces, such as "shop", "forge", "manufactory" and "factory" in

³⁶ For a discussion of the urban character of industrial activity in Ontario in 1871 (including an explanation of places defined as "proto-urban") see Elizabeth Bloomfield and G.T. Bloomfield. The Ontario Urban System at the Onset of the Industrial Era, #3 in this series of research reports.

English and "atelier", "boutique", "fabrique" or "manufacture" in French, were considered possibly significant in contemporary perceptions of industrial settings. Regional variations in industrial language of all kinds were also thought to be interesting.³⁷ Third, we were making such judgements in assigning Standard Industrial Classification codes to every record, so there was no need to standardize the terms for data processing purposes.³⁸

Transcribers thus included words like "shop", "forge", "factory", "manufactory" and "works" if there was space available. English examples included:

Shoemaker Shop
Blacksmith Forge
Boot/Shoe Factory
Marble/Stone Works

Such terms for the various work environments could be abbreviated if space required. An establishment combining two or more products or services had both keywords entered and linked by a / , and the entry probably had to be abbreviated as well. This abbreviation could take several forms: a pluralization of the second product or service; a shortened form of the words shop, forge, factory, mill, or works; or an abbreviated form of one or both of the main products or services. For example:

Blacksmith/Carriages
Wagons/Blacksmith
Tannery/Boots/Shoes
Saw/Shingle/Grist M (for Mill)
Carriage/Sash/Door F (for Factory)
Brass Fndry/Lamp Fcy
Moulding/Planing Fcy
Blcksmith/Machine Sp (for Shop)
Wagon/Carriage Shop

In cases where the census enumerators added an "s" to certain types of establishments indicative of the possessive, these were transcribed without the "s": Thus "Carpenters Shop" became "Carpenter Shop", and "Tailors Shop" became "Tailor Shop". Where enumerators may have used compound nouns to describe the type of establishment, such as "Shoe Maker", "Cabinet Maker" or "Carriage Maker", these terms were contracted into single words for consistency and brevity. Thus "Cabinet Maker" became "Cabinetmaker" and so on. If the enumerator had entered two establishments with separate numeric data under one proprietor's name, they were be entered individually as separate establishments. Often the floating capital was the only joint figure

³⁷ This consideration led us to compile lists of terms used in the census to describe types of establishment, kinds of raw materials and products and units of measurement. See the Glossary of Industrial Language, #5 in this series, and French-English Dictionary of Industrial Language, #6 in this series.

³⁸ See later section and Appendix A-2 in this report and for more detail, Standard Industrial Classifications Applied to Historical Data: the Case of the 1871 Industrial Census, #7 in this series of research reports.

provided and this would be allocated between the two establishments in proportion to the value of production of each.

French expressions for kinds of establishment were also slightly modified in data entry. The enumerators in French-language districts usually began with a generic word such as "atelier", "boutique", "manufacture", "moulin", or "fabrique" to describe the type of workplace and followed it with the particular activity. "Atelier" and "boutique" connoted a smaller scale of operation and the use of manual rather than inanimate power, "atelier" being used for the workplaces of artists, dressmakers, florists, cabinetmakers, printers, milliners, pastrycooks, pharmacists, photographers and upholsterers and "boutique" for bakers, hatters, carpenters, wheelwrights, shoemakers, tinsmiths, blacksmiths, goldsmiths, bakers, saddlers, tailors, tanners and coopers. "Moulin" was consistently used for the types of powered work settings where "mill" would be used in English, such as a saw mill, flour mill or carding mill, though a woolen textile mill was described as a "Fabrique de drap". The terms "manufacture" and "fabrique" were apparently used for larger scale enterprises than the "ateliers" and boutiques", roughly equivalent to manufactories and factories.

To fit the details in our available space and to give prominence to the particular type of industry, we entered this keyword first and followed it with an abbreviation for the generic terms that described the type of workplace. Thus "Atelier de cordonnier" was entered as "Cordonnier, A"; "Boutique de ferblantier" as "Ferblantier, B"; "Moulin à scie" as "Scie, MI"; "Manufacture de chassis" as "Chassis, Mf"; "Fabrique de meuble" as "Meuble, F"; and "Machine à bardeau" as "Bardeau, Mch".

Typeword and Force The type of power was entered in full and, an exception to the general rule of transcribing, the English form was always entered as this variable was used for analysis of types of power from an early stage of the project.³⁹ The main types of power thus were WATER, STEAM, or HORSE. If more than one kind of power was given in the manuscript schedules, each was abbreviated to its initial letter. So "Water and Steam" was entered as "W/S". In the case of establishments using horse power, the word "HORSE" would be entered for type of power but the FORCE space was occasionally left blank. In such cases, the numeral "1" would be entered for FORCE.

Kinds, quantities and values of Raw Materials or Products

The 1871 census schedules allowed space for enumerators to complete details of the quantities as well as the dollar values of raw materials and of manufactured products. But the census organizers anticipated problems with the returns for these measures, stating in the "Instructions to Officers" that "in many instances the raw materials or articles manufactured are of such a multifarious character that they must be lumped together and entered by the

³⁹ See, for example, G.T. Bloomfield and Elizabeth Bloomfield, Water Wheels and Steam Engines: Powered Establishments of Ontario, #2 in this series of research reports.

value".⁴⁰ Significantly, the columns for values of raw materials and values of products are headed "Aggregate value".

There are several problems with these quantity data. The census staff differed in their handling of this part of the schedule, some making considerable efforts to ascertain and record the types, quantities and values of component raw materials and manufactured products, and setting these out systematically and clearly. Other simply named one or several materials or products but did not specify separate quantities or values.

We did not code or enter these quantity data in the first stage of the project, partly because of the incompleteness and variability from one census enumerator to another. We were also concerned that the extreme variety of materials and products and of units of measurement -- feet, bushels, tons, pounds weight, not to mention quintals and toises -- would have made the database too large and unwieldy, as at that time we were constrained by the format of 80-column coding cards. We still do not know how much significance can be extracted from these quantity data, when census enumerators varied so considerably in their practices. However, in keeping with our goal of providing as nearly as possible a facsimile record of the manuscript census, we have now entered all such details as were given for quantities of raw materials and manufactured products, as well as units of measurement when available.

In the vast majority of records, details for raw materials could be fitted into one set of fields in the computer record, as could details for products. Only in 4.5 per cent of Maritime establishments, 4.9 per cent of those in Quebec and 6.5 per cent of those in Ontario, were more than two sets of fields required for products or raw materials. The quality of the data for products and raw materials is assessed in a later section of this report.

Great care was taken in the transcription of the information for raw materials and products, as the following examples will show. If separate quantities and/or values of raw materials and products were recorded in the manuscript schedules, these were entered separately in the computer record. The examples given below show how the data would be set out on Lotus spreadsheets:

Raw Material Examples:

KIND	UNIT	QUANTITY	VALUE
1. LEATHER,UPPER	SIDE	24	84
LEATHER,SOLE	SIDE	12	72
SKINS,CALF		12	36
2. LUMBER	FT	300000	6000
HARDWARE			1000
3. LOGS,SPRUCE,PINE,BIRCH	FT BM	2000000	12000

⁴⁰ "Manual", Canada Sessional Papers (1871): 139.

Product Examples:

KIND	UNIT	QUANTITY	VALUE
1. BOOTS	PR	700	2800
SHOES	PR	300	1200
2. WAGONS		5	480
3. PLANKS	FT BM	55000	390

If only a list of raw materials or products were given and only one aggregate value and quantity, they were entered in a string separated by one or more slash marks:

Raw Material Examples:

KIND	UNIT	QUANTITY	VALUE
1. LUMBER/GLASS/PAINT/HARDWARE			50000
2. BRASS/TIN/LEAD/COPPER	Q	208	8000

Product Examples:

KIND	UNIT	QUANTITY	VALUE
1. CARRIAGES/SLEIGHS/WAGONS		60	10000
2. LOCKS/PIPES,BRASS/BOILERS			75000

If separate quantities were given for each raw material or product, but only one value, this aggregate value was entered separately so that it would not be confused for a specific value for an individual raw material or product. At times it was not obvious whether a quantity or value referred specifically to one or both of the raw materials or products; it was not be assumed that it did unless the information was absolutely clear. Such a problem is illustrated in the following:

Raw Material Examples: (as set out on the manuscript schedule)

KIND	UNIT	QUANTITY	VALUE
1. TIMBER/LUMBER	FT BM	60000	7000
IRON/COPPER	TON	160	
2. LUMBER,SPRUCE,BIRCH			1200
IRON	TON	6	

The value figure given could not be assumed to refer to just the TIMBER/LUMBER in the first example, nor just to the LUMBER in the second example, since an overall aggregate value of raw materials is given. So they would be entered as follows:

Raw Materials: (the above as entered on Lotus spreadsheet)

KIND	UNIT	QUANTITY	VALUE
1. TIMBER/LUMBER	FT BM	60000	
IRON/COPPER	TON	160	
			\$ 7000
			(separate aggregate value)
2. LUMBER,SPRUCE,BIRCH			
IRON	TON	6	
			\$ 1200
			(separate aggregate value)

The same general practice is illustrated by the following examples of products:

Product Examples: (as in the manuscript schedules)

KIND	UNIT	QUANTITY	VALUE
1. CARRIAGES			320
REPAIRS			
2. CARRIAGES/SLEIGHS	NO	6	700
REPAIRS			

The value of the products given in these examples cannot be assumed to refer to just CARRIAGES or to CARRIAGES/SLEIGHS. So they would be entered:

Products: (the above as entered on Lotus spreadsheet)

KIND	UNIT	QUANTITY	VALUE
1. CARRIAGES/REPAIRS			320
2. CARRIAGES/SLEIGHS/REPAIRS	NO	6	700

In some cases, the statement of kind of raw materials or products was followed by several qualifiers, as in these examples:

LOGS,SPRUCE,PINE,BIRCH,HEMLOCK
GRAIN,WHEAT,OAT,RYE,OTHER

Lengthy phrases such as "of different kinds" or "of various kinds" were usually replaced by the word "ASSORTED".

Units of measurement Quite often, units of measurement for raw materials and products were not stated but could be inferred; however, the transcriber did not enter such inferences or assumptions. Moreover, units were transcribed separately in English and French as stated by the original enumerators. No attempt was made to give standardize equivalent measures which made different names in English and French, such as "hundredweight" or "cwt" in English and "quintals" in French. Consistent abbreviations were used for units of measurement where they were longer than six characters (Appendix A-4).

Comments Information appearing on the Remarks/Remarques column of the schedule itself was entered in the COMMENTS field, together with additional information from any other column. Two hundred character spaces were allowed in the record structure for this purpose. Most commonly such other information related to the proprietor and type of establishment. A few zealous enumerators entered more precise locational information in the Remarks column, such as street address in urban places or locality name in some rural districts. General comments relating to the whole census enumerator's division were sometimes written on the last of the manuscript schedules for that particular division. Overall, 14 per cent of the establishments in New Brunswick and Nova Scotia have some details in the COMMENTS fields, and 21 per cent of those in Ontario and Quebec.

Industrial code A Standard Industrial Classification (SIC) code was added for each establishment record. The SIC code assigned to each establishment is vital for retrieval and analysis of records in the same or related lines or business. The 1871 census organizers attempted no logical organization of industrial types; and not until the 1920s were Canadian census reports organized by anything other than an alphabetical list of industrial products.⁴¹ Understandably, there were some problems in applying a 1970 classification to the conditions of a century earlier. The more rudimentary level of economic organization in 1871 meant that many enterprises combined industrial activities that would more typically be separated in specialized businesses by 1970. In addition, the repairing of all kinds of goods and equipment was more integrally associated with manufacture in 1871.

The 1970 classification system was adapted to the conditions of 1871 in two main ways. Suffixes of hyphen and capital letter were added to the basic 3-digit code to give greater specificity, a brewery (109-B) being distinguished from a distillery (109-D), for example. An establishment which combined two or more products or services was designated with SIC codes of the two most important linked by a slash (/), so that a flour mill-cum-distillery is represented as 105/109-D. Appendix A-2 lists the more common SIC codes (those used at least 20 times) that were assigned in making the 1871 industrial census data machine-readable. In addition, common compound codes used at least 40 times to describe enterprises that combined two or more principal types of products were:

- 105/105-O (flour and oatmeal mill),
- 107/108-C (bakery and confectionery),
- 179-S/179-T (saddlery/trunkmaking),
- 244/249-M (dressmaking and millinery),
- 249-M/244 (millinery and dressmaking),
- 251-S/259-C (shinglemaking and cooperage),
- 251/251-S (saw mill and shinglemaking),
- 254/261 (sash, door and blind factory combined with furniture and cabinet making),

⁴¹ Dominion Bureau of Statistics, Standard Industrial Classification Manual (revised edition, 1970). For a full discussion of our application of this system to the 1871 census, see Standard Industrial Classifications Applied to Historical Data: the Case of the 1871 Industrial Census, #7 in this series of research reports.

261/258 (cabinet and coffin making),
 261/329 (cabinet and carriage making),
 307-S/311 (manufacture of stoves and agricultural implements)
 311/329 (agricultural implement and carriage making)
 329/896 (carriage making and blacksmithing).

The appropriate code was determined by study of both the stated kind of establishment and the actual products reported in each case. "Type of establishment" information alone could be quite misleading, as this term was sometimes an occupational label which did not really describe the actual products. Thus a woman might be described as a "milliner" but if the products of her business included only dresses and other women's clothing rather than bonnets, hats and women's head-dresses, she would be assigned a code of 244 for women's clothing and not 249-M for millinery. Or a business might be described very generally as a "foundry" which would rate a code of 294 if its products were primarily iron castings. But more typically its products were fabricated metal products such as stoves (307-S) or machinery such as agricultural implements (311). Or one who was by trade a blacksmith and so recorded in the kind of establishment may have had opportunities to make all kinds of metal goods, often including carriages and wagons. If such a person made definite numbers of carriages or other metal products, his business would be assigned the appropriate and more specific SIC code of 329 rather than "896" which was left for a blacksmith engaged only in "country work and repairs".

The assigning of the Standard Industrial code was clearly a very different task from the transcription work in most other fields of the database. All but the most experienced staff found it easier to assign SIC codes as a separate phase after the basic data had been entered. The SIC codes first assigned by the staff were later subjected to computer-assisted global editing in which the whole database or major parts of it were sorted or indexed on SIC code. By examining the coding of records in their SIC context, anomalies could be detected quite easily. The use of the SIC code is discussed in more detail in another research report in this series.

5 HOW GOOD ARE THE 1871 MANUSCRIPT CENSUS DATA?

How good were the data of industrial activity in 1871? The census organizers declared that they were "as accurate as is humanly possible", but questions inevitably arise. How complete was the coverage of industrial establishments operating in 1870-1? How thoroughly did enumerators question proprietors on aspects of their industrial activities and how carefully did they records these on the census schedules? How good is the match between the totals we obtain by computerized aggregations of the data from the manuscript schedules and the tabulations published in the 1870s?

No administrative reports or critical appraisals of the 1871 Census of Canada at the time have survived, so we do not how those responsible at the time regarded the quality of the data. We trust that the assessment by the commissioner for Pictou County, Nova Scotia was atypical; he remarked that "the returns secured by enumerators from industrial establishments in this

county are for the most valueless."⁴² The report on the Ninth Census of the United States in 1870 did contain some observations on problems of definitions and enumeration procedures which seem to have been similar to those in Canada. The authors of the U.S. report recommended that the industrial census should be taken by those with some technical expertise in industrial processes, that only specifically industrial activities should be counted and such businesses as carpenters, blacksmiths, coopers, painters, plasterers and plumbers should be excluded. They also emphasized the "manifest uselessness and indeed, impracticability, of returning the kinds and quantities of materials in a form allowing of systematic tabulation".⁴³

In this preliminary evaluation of the industrial data of the 1871 Census of Canada, we comment first on problems of missing data for individual establishments in the manuscript schedules, then on the incidence of missing records, and finally on the discrepancies between the published totals and those obtained by computer aggregations of the manuscript data.

Missing data

A fair proportion of the establishment records have missing data. Table 13 presents a summary of the percentages of missing data for the different variables in each of the three major regions, Ontario, Quebec and the two Maritime provinces. Records with missing data form less than one per cent of all records for the proprietor's name, value of production and number of employees (slightly exceeded in the Maritimes). Percentages missing are in the range of one to five per cent for number of working months in the year, value of raw materials, wages (slightly exceeded in Ontario) and number of horse-power units in steam-powered establishments (slightly exceeded in the Maritimes).

Value of fixed capital is generally missing in between five and ten per cent of all records (slightly better in Quebec). Figures for floating capital and for horse-power units in water-powered establishments are missing in significantly more cases generally, and in over one third of Maritime records for each variable. The higher percentages of missing data for floating capital and for horse power units in water-powered businesses may explain why, in the published reports of the 1871 census, capital means fixed capital only and also why no data on the use of power in 1871 were ever published.

Percentages of missing data for the derived variable of VADD (value added in manufacturing) reflect the percentages of missing data for the variables of SUMRAWC and SUMPROC that are used in calculating these measures. Similarly AVWAGE (the average monthly wage per employee) reflects the quality of data for the three basic variables of TOTEMP, WAGES and MONTH.

⁴² Remark in COMMENTS field of first record in CD 200: Pictou County; CSD J: Gairloch.

⁴³ Superintendent of Census, The Statistics of the Wealth and Industry of the United States compiled from original returns of the Ninth Census (Washington: Government Printing Office, 1872), pp. 383-5.

If one expresses the percentages in Table 13 in positive terms, it is possible to state the percentages of "complete" records in which basic numeric data are not missing. Records in which data are not missing for any of the following variables -- employment, fixed capital, months, wages, value of raw materials, value of products or (in the case of powered firms) units of horse power -- constitute 84 per cent of all in Ontario, 86 per cent of those in Quebec and 76 per cent of those in the Maritimes.

Table 13
Missing data in CANIND71 database (percentages by region)

Variable	Ontario	Quebec	Maritimes
PROPRIOR	0.5	0.4	0.5
MONTH	0.9	1.9	4.4
FIXCAP \$	5.6	4.6	8.5
FLOCAP \$	22.5	12.2	35.7
FORCE if TYPEPOW=WATER	11.8	17.5	38.7
FORCE if TYPEPOW=STEAM	1.1	1.2	6.0
TOTEMP	0.5	0.5	1.2
WAGES \$	6.0	4.2	1.8
SUMRAWC \$	3.5	1.8	2.6
SUMPROC \$	0.9	0.4	0.8
AVWAGE (derived)	8.8	6.4	6.6
VADD (derived)	4.5	2.5	3.2
Composite (non-power)	17.7	12.0	17.5
Composite (powered)	11.7	18.4	36.8
COMPOSITE (merged)	16.1	13.6	24.0

Note: Composite measure of missing data based on variables of FIXCAP, TOTEMP, MONTH, WAGES, SUMRAWC, SUMPROC and (if establishment reported power) FORCE.

We may guess at the reasons for missing data in some cases. Presumably most of the cases in which no employees were reported were small artisanal craftshops, in which the proprietor did not reckon himself or herself to be an employee. The same may be true for the small number of establishments (0.2 per cent in Ontario and the Maritimes and 0.4 per cent in Quebec) in which only boys or girls were reported as employees but which must really have had adult proprietors. For the purposes of analysis, some adjustments might be made for these missing data but we have not interpolated them.

In some of the records in which there are no entries for the aggregate cost of raw materials, this can be interpreted as referring to custom work in which the clients brought the materials to the shop to be processed or made up. Strictly, enumerators were instructed to enter the real value of the raw materials or products, regardless of whether these amounts passed through the accounts of the proprietor.⁴⁴ However, in practice it must have been difficult to get information from proprietors of grist and carding mills who worked for a toll or proportion of the grain or wool that they processed. Similarly, tailors and dressmakers used fabric and other material provided by their customers. The enumerators sometimes refer, in the Remarks column, to the practice of custom work and the difficulties of obtaining accurate figures for the value of raw materials and products. In such cases, for the purposes of statistical analysis, the value of production entered might be interpreted rather as a measure of added value. The data for our derived variable of VADD is naturally affected by under-reporting of values for raw materials and products, especially for businesses engaged in custom work. The value of VADD (calculated only for records in which neither SUMRAWC nor SUMPROC was missing) is negative in a significant number of such cases.

If enumerators entered at least aggregate values for raw materials and products in the vast majority of cases, only in a small proportion of records did they provide considerable detail of inputs and outputs. Cases such as John Watson's foundry in Ayr, southern Waterloo County, are exceptional in detailing so many raw materials and/or products in both quantities and values (Figure 6). More characteristic is the entry for Robert Hay and Company of Toronto, by far the largest business making cabinets and upholstery in Ontario with production valued at \$500,000 in 1871. This record specified for materials only "140,000 ft lumber" and for products "cabinet ware of all kinds". Altogether, less than one per cent of all the records in the CANIND71 database listed at least three products and three raw materials together with some details of quantities and component values. For only 35 establishments in all four provinces are at least five raw materials and five products listed with their quantities in the 1871 manuscript census schedules.

How good are the details of component raw materials and products in different districts and localities? To test an impression that quantities as well as values of raw materials and products were more often stated for simple industrial activities in rural areas, we examined two contrasting districts in the province of Quebec. The West Ward of the Montréal Centre census district and Mégantic census district each had about 195 establishments in total (Table

⁴⁴ See page 4 and footnote 14, above.

14). The Montréal ward recorded 6,266 employees producing \$7.7 million worth of a great variety of industrial types, while Mégantic had only 390 employees making under \$400,000 worth of typically rural products in small-scale establishments. Only six firms in Mégantic had an output worth at least \$10,000 while Montréal's West Ward had 102 businesses reporting output of at least \$10,000, nineteen of which were complex enterprises producing over \$100,000 worth of output.

The Montréal enumerators, whose work was otherwise painstakingly accurate and detailed, listed kinds of raw materials or products in the form of strings of commodities without any details of quantities or component values for four of every five businesses (type E in Table 14). The only numeric data were aggregate dollar values for all raw materials and for all products of each business. Only in ten records (type A) were both inputs and outputs specified in a fair degree of detail, and in a further seventeen (type B), inputs only were detailed. In Mégantic, on the other hand, two-thirds of all the records had some details of the quantities as well as values of raw materials and products (type A), and a further 24 per cent provided such details for inputs only (type B). However, virtually all the Mégantic records describe very simple industrial activities in which there was only one main raw material and one product. A typical entry, 300 cords of pinewood worth \$900 to make one million shingles worth \$2000, could provide specific details of quantity and unit of measurement as well as value and still fit easily on one line of the schedule (and thus in one set of fields in the record structure). Only seventeen of Mégantic's records use a second line or set of fields for either raw materials or products.

Table 14
Assessment of raw material and product data in two Quebec districts, 1871
percentages

	Montréal Centre West Ward	Mégantic
Type A: Inputs & outputs	5.2	67.3
Type B: Inputs only	8.8	24.4
Type C: Outputs only	3.0	0.5
Type D: Partial details	1.5	-
Type E: Lists of materials & aggregate value only	80.0	5.1
Type F: No materials specified & aggregate value	1.5	2.5
TOTALS	100.0	100.0

Missing records

Some industrial businesses, definitely operating in 1871, seem to have been missed by the census enumerators. While 27 gasworks were included in the census record, only three waterworks systems were counted throughout Canada. One of those missed was the Hamilton waterworks which was located outside the city in Saltfleet Township. Another example of a missing record is for the Wisner agricultural implements business in Brantford. Some enterprises seem to have been missed, as with the Cornwall Manufacturing Company's woolen mill, until one realizes that John Warwick the manager who presumably provided the enumerator with details of the enterprise is stated in error to be the proprietor.

Any comparison of the enterprises listed in the 1871 manuscript census for a particular place with entries in a directory or the R.G. Dun reference books for the same year suggests that some smaller enterprises may have been missed.⁴⁵ There are always some establishments that have been listed in the industrial census that one cannot find in directories or Dun reference books, and conversely some businesses in Dun and the directories that are missing from the census.

Another category of missing records consists of industrial schedules that have been lost since 1871. In a few CEDs or whole CSDs, businesses were enumerated in 1871 and their data were included in the census tabulations published in the 1870s, but the manuscript schedules were later lost. The most serious loss is of all the industrial data for King's Ward in the city of Saint John. A comparison of the published statistics and of the incomplete manuscript data shows that at least 210 establishments, 1100 employees, \$250,000 value of fixed capital and \$1,350,000 worth of output are missing because of the loss of the schedules for King's Ward. In Nova Scotia, industrial schedules were lost for Kentville and parts of Centreville and Somerset CSDs in King's County, three CSDs of Halifax West and, more seriously, for most of the CSDs of Shelburne County. The loss of the Shelburne County schedules means that at least 75 establishments, 250 employees, \$65,000 worth of fixed capital and \$235,000 worth of output are missing. In comparison, the incidence of missing records in Ontario and Quebec is negligible.

Differences between published census and manuscript totals

We have found that the 1871 published totals for census districts may differ considerably from those we have obtained from our computerized sums of the individual establishment data. Of course, the missing schedules in New Brunswick and Nova Scotia affect the aggregations for those provinces and any comparisons between published and manuscript totals. Otherwise, it seems that the published figures usually understate the real totals obtained by aggregations of the individual establishment data.

⁴⁵ Undercounting was most common with the smaller artisanal businesses in the U.S. manuscript census as well, according to John B. Jentz, "A Note on Evaluating the Error in the Gilded Age Manufacturing Census: The Problem of the Hand Trades", Historical Methods Newsletter 15 (1982): 79-81.

Table 15 shows the variance between the manuscript and published data for each of the four provinces and Canada as a whole. A plus sign indicates that the manuscript data exceed the published data by the percentage margin stated for each variable. A minus sign means that the manuscript data fall short of the published data by the stated percentage. Table 16 presents the equivalent data for the census districts containing the six largest Canadian cities in 1871, Montreal, Quebec City, Toronto, Hamilton, Saint John and Halifax as well as Wellington County and Simcoe County in Ontario.

Table 15
Percentage variance between manuscript and published census totals, 1871

Variable	Ontario	Quebec	New Brunswick	Nova Scotia	CANADA
Establishments	+ 11.4	+ 4.3	+ 15.3	- 0.7	+ 9.0
Fixed capital	+ 22.5	+ 5.9	- 9.8	+ 1.3	+ 12.5
Men employed	+ 6.5	+ 3.8	- 6.3	+ 1.6	+ 3.9
Women employed	+ 20.1	+ 3.9	+ 5.0	- 12.1	+ 10.3
Boys employed	+ 6.7	+ 3.8	- 4.2	+ 3.8	+ 4.1
Girls employed	+ 12.5	+ 7.8	+ 8.2	- 3.4	+ 8.8
Total employed	+ 9.2	+ 4.1	- 5.2	+ 0.8	+ 5.3
Wages	+ 5.5	+ 4.2	- 11.5	+ 4.6	+ 3.5
Raw materials	+ 2.3	+ 5.4	- 11.3	- 0.1	+ 2.2
Products	+ 4.5	+ 3.7	- 12.9	+ 4.6	+ 2.9

Source: Manuscript census data compiled from CANIND71 database. Published 1871 Census Volume III (Table 54) for fixed capital in dollars, total employed, wages in dollars, raw materials in dollars, and products in dollars. Numbers of establishments and of men, women, boys and girls employed from the machine-readable version of the published data for individual industrial types in 1871 (Tables 28 to 53).

Table 16
Percentage variance between manuscript and published totals, 1871 census

	Montreal	Quebec	Toronto	Hamilton	St John	Halifax	Wellington	Simcoe
census districts	104/105/106	145/146/147	46/47	24	174	196	33/34/35	41/42
Establishments	+ 4.2	+ 2.8	+ 8.6	+ 6.8	- 20.9	- 5.6	+ 11.5	+ 26.1
Fixed capital	+ 5.3	+ 3.4	+ 173	+ 2.3	- 26.0	- 1.9	+ 6.0	+ 4.1
Men employed	+53.9	+ 0.7	+18.2	+43.3	- 33.0	+ 4.7	+ 13.8	+ 4.7
Women	- 2.1	+ 2.6	+ 1.7	+ 0.1	- 39.1	-31.2	+ 24.2	+153.1
Boys	+ 1.7	+ 3.7	+ 2.6	+ 1.9	- 31.9	- 4.6	+ 13.2	+ 8.9
Girls	- 0.4	+ 7.7	+ 3.5	- 4.1	- 24.8	-22.1	...	- 53.0
Total employed	+ 4.6	+ 1.6	+12.6	+29.6	- 33.8	- 0.7	+ 14.7	+ 9.1
Wages	+ 7.7	- 0.4	+12.7	+41.9	- 32.2	+ 9.5	+ 10.6	+ 4.2
Raw materials	+ 1.8	+ 1.5	+ 5.0	+ 1.7	- 30.4	- 4.8	+ 4.3	+ 3.1
Products	+ 4.4	+ 1.6	+12.6	+ 3.5	- 32.3	+ 5.1	+ 6.4	+ 8.7

Source: Manuscript census data compiled from CANIND71 database. Published 1871 Census Volume III (Table 54) for fixed capital in dollars, total employed, wages in dollars, raw materials in dollars, and products in dollars. Numbers of establishments and of men, women, boys and girls employed from the machine-readable version of the published data for individual industrial types in 1871 (Tables 28 to 53).

The variance between published industrial census totals and those obtained by computer aggregations of the manuscript data poses methodological problems. In this preliminary probe, we have noted a variance of up to 22 per cent in the totals for employees, wages, raw materials and value of output for whole provinces. For individual census districts the percentage variance is much higher (Table 16). The published totals understate those derived from the manuscript data in almost all cases. The percentage of understatement in the published data is apparently highest in Ontario, with 11 per cent more establishments, 20 per cent more women employed in industry and 22 per cent higher value of fixed capital in the manuscript data than the published tables. Quebec's totals vary less widely between the manuscript and the published sources. In New Brunswick and Nova Scotia, the incidence of missing schedules is a special factor that is reflected in the large percentage differences for those provinces (Table 15) and for Saint John and Halifax census districts (Table 16).

Can these disparities be explained simply as arithmetical error by those who added the statistics from the manuscript schedules in the early 1870s? Some errors of this kind can be detected by recalculating a machine-readable version of the published tabulations.

Did the Ottawa clerks who compiled the manuscript returns from the districts into the published tables also systematically edit the manuscript returns to exclude certain types and sizes of industrial establishments? It is suggestive that for all of Ontario, 95 per cent of the male industrial workers counted from the manuscript schedules were reported in the published census, but only under 85 per cent of the women. A few industry types in which women were prominent, such as handloom weaving, seem to have been systematically excluded from the published tabulations. A microscale study linked with our project has found that substantial numbers of female weavers in one district of eastern Ontario, whose operations were fully recorded in the manuscript industrial census, were totally excluded from the published census returns.⁴⁶ Women would seem to have been included in the published data more consistently when they worked in larger establishments such as factories in urban centres than when they worked more informally in smaller workplaces in rural settings, including workshops in their own homes. The percentage variance for women is much lower in the larger urban centres than in the counties such as Wellington and Simcoe which were mainly rural (Table 16).

Railway workshops, of which the largest were those of the Grand Trunk in Montreal and Brantford and the Great Western in Hamilton, were counted in the manuscript schedules but dropped from the published tabulations.⁴⁷ This

⁴⁶ Janine Grant and Kris Inwood, "How Urban was Cloth Manufacturing in 1870?", paper prepared for joint session of Canadian Economics Association and Canadian Historical Association, Hamilton, June 5 1987. A revised version of this paper will be published in Canadian Papers in Rural History (1990).

⁴⁷ Noted also by Craven and Traves, who caution that the aggregate tables in the published reports for the 1871 census should be used with the greatest caution for this industry group. See "Canadian Railways as

exclusion is reflected in the percentage variance for male employees in Montreal and Hamilton, for fixed capital in Toronto and for wages in Hamilton and Toronto (Table 16).

The discovery of such variance between the manuscript and published industrial data of the 1871 Census raises some questions about the findings of Chambers and Bertram, Gilmour and Bland, who based their research on the published census material.⁴⁸ The reliability of the published U.S. censuses of manufacturing in the nineteenth century has also been questioned.⁴⁹ It also enhances the significance of the original manuscript schedules and indicates the need to code and process data for all establishments, not just those in a sample.

Differences between tabulations derived from the manuscript schedules and the tabulations published in the 1870s might well be analyzed more thoroughly. The general questions of the reasons for the calculated differences would be illuminated by analysis of the percentage variance in particular geographical districts and industry types.

Manufacturers, 1850-1880," p. 264.

⁴⁸ See note 16 above and W.R. Bland, "The Changing Locational Pattern of Manufacturing in Southern Ontario from 1881 to 1932," Ph.D. Thesis, Indiana University, 1970. Discrepancies between the published and manuscript totals were noted by Kealey in Toronto Workers, Appendix 1.

⁴⁹ Margaret Walsh, "The Value of Mid-Nineteenth-Century Manufacturing Returns: The Printed Census and the Manuscript Census Compilations Compared", Historical Methods Newsletter 4 (1971): 43-51. Bateman and Weiss found that the published summaries were "frequently inaccurate and in a few cases contain gross inaccuracies and omissions" and concluded that their small random samples "provide a better description of the parent population than do the published census summaries" (A Deplorable Scarcity pp.169-171).

Appendix A-1:

CENSUS DISTRICTS IN 1871

CDID CENSUS DISTRICT NAME

ONTARIO

001 ESSEX
 002 KENT
 003 BOTHWELL
 004 LAMBTON
 005 ELGIN WEST
 006 ELGIN EAST
 007 MIDDLESEX WEST
 008 MIDDLESEX NORTH
 009 MIDDLESEX EAST
 010 LONDON
 011 NORFOLK SOUTH
 012 NORFOLK NORTH
 013 OXFORD SOUTH
 014 OXFORD NORTH
 015 BRANT SOUTH
 016 BRANT NORTH
 017 HALDIMAND
 018 MONCK
 019 WELLAND
 020 NIAGARA
 021 LINCOLN
 022 WENTWORTH SOUTH
 023 WENTWORTH NORTH
 024 HAMILTON
 025 HURON SOUTH
 026 HURON NORTH
 027 BRUCE SOUTH
 028 BRUCE NORTH
 029 PERTH SOUTH
 030 PERTH NORTH
 031 WATERLOO SOUTH
 032 WATERLOO NORTH
 033 WELLINGTON SOUTH
 034 WELLINGTON CENTRE
 035 WELLINGTON NORTH
 036 GREY SOUTH
 037 GREY NORTH
 038 HALTON
 039 PEEL
 040 CARDWELL
 041 SIMCOE SOUTH
 042 SIMCOE NORTH
 043 YORK NORTH
 044 YORK WEST
 045 YORK EAST
 046 TORONTO WEST
 047 TORONTO EAST
 048 ONTARIO SOUTH
 049 ONTARIO NORTH
 050 DURHAM WEST
 051 DURHAM EAST
 052 VICTORIA SOUTH
 053 VICTORIA NORTH

CDID CENSUS DISTRICT NAME

054 NORTHUMBERLAND WEST
 055 NORTHUMBERLAND EAST
 056 PETERBOROUGH WEST
 057 PETERBOROUGH EAST
 058 PETERBOROUGH NORTH
 059 PRINCE EDWARD
 060 HASTINGS WEST
 061 HASTINGS EAST
 062 HASTINGS NORTH
 063 LENNOX
 064 ADDINGTON
 065 FRONTENAC
 066 KINGSTON
 067 LEEDS SOUTH
 068 BROCKVILLE
 069 GRENVILLE SOUTH
 070 LEEDS N/GRENVILLE
 071 DUNDAS
 072 STORMONT
 073 CORNWALL
 074 GLENGARRY
 075 PRESCOTT
 076 RUSSELL
 077 OTTAWA
 078 CARLETON
 079 LANARK SOUTH
 080 LANARK NORTH
 081 RENFREW SOUTH
 082 RENFREW NORTH
 083 NIPISSING SOUTH
 084 NIPISSING NORTH
 085 MUSKOKA
 086 PARRY SOUND
 087 MANITOULIN
 088 ALGOMA EAST
 089 ALGOMA CENTRE
 090 ALGOMA WEST

QUEBEC

091 PONTIAC SOUTH
 092 PONTIAC NORTH
 093 OTTAWA WEST
 094 OTTAWA CENTRE
 095 OTTAWA EAST
 096 ARGENTEUIL
 097 DEUX MONTAGNES
 098 LAVAL
 099 TERREBONNE
 100 L'ASSOMPTION
 101 MONTCALM
 102 JOLIETTE
 103 BERTHIER
 104 MONTREAL CENTRE
 105 MONTREAL EST
 106 MONTREAL WEST

CDID	CENSUS DISTRICT NAME	CDID	CENSUS DISTRICT NAME
107	HOCHELAGA	160	DORCHESTER EST
108	JACQUES CARTIER	161	BELLECHASSE NORD
109	VAUDREUIL	162	BELLECHASSE SUD
110	SOULANGES	163	MONTMAGNY
111	BEAUHARNOIS	164	L'ISLET
112	CHATEAUGUAY	165	KAMOURASKA
113	HUNTINGDON EAST	166	TEMISCOUATA
114	HUNTINGDON WEST	167	RIMOUSKI OUEST
115	LAPRAIRIE	168	RIMOUSKI EST
116	NAPIERVILLE	169	BONAVENTURE
117	ST-JEAN	170	GASPE OUEST
118	CHAMBLY	171	GASPE CENTRE
119	VERCHERES	172	GASPE SUD
120	RICHELIEU	173	ILES DE LA MADELEINE
121	ST-HYACINTHE		
122	BAGOT		
123	ROUVILLE		<u>NEW BRUNSWICK</u>
124	IBERVILLE	174	ST JOHN
125	MISSIQUOI	175	CHARLOTTE
126	BROME	176	KING'S
127	SHEFFORD	177	QUEEN'S
128	MASKINONGE	178	SUNBURY
129	ST-MAURICE SUD	179	YORK
130	ST MAURICE NORD	180	CARLETON
131	TROIS-RIVIERES	181	VICTORIA
132	CHAMPLAIN SUD	182	RESTIGOUCHE
133	CHAMPLAIN NORD	183	GLOUCESTER
134	YAMASKA	184	NORTHUMBERLAND
135	NICOLET	185	KENT
136	DRUMMOND	186	WESTMORELAND
137	ARTHABASKA	187	ALBERT
138	RICHMOND		
139	WOLFE		<u>NOVA SCOTIA</u>
140	SHERBROOKE		
141	STANSTEAD	188	HANTS
142	COMPTON	189	KING'S
143	PORTNEUF	190	ANNAPOLIS
144	QUEBEC COMTE	191	DIGBY
145	QUEBEC OUEST	192	YARMOUTH
146	QUEBEC CENTRE	193	SHELBURNE
147	QUEBEC EST	194	QUEEN'S
148	MONTMORENCY	195	LUNenburg
149	CHARLEVOIX	196	HALIFAX, WEST
150	CHICOUTIMI	197	HALIFAX, EAST
151	SAGUENAY	198	CUMBERLAND
152	LABRADOR	199	COLCHESTER
153	LEVIS VILLE	200	PICTOU
154	LEVIS COMTE	201	ANTIGONISH
155	LOTBINIERE	202	GUYSBOROUGH
156	MEGANTIC	203	INVERNESS
157	BEAUCE OUEST	204	VICTORIA
158	BEAUCE EST	205	CAPE BRETON
159	DORCHESTER OUEST	206	RICHMOND

Appendix A-2

Standard Industrial Classification (1970) adapted for 1871:
short list of types with at least 20 occurrences in CANIND71 database

DIVISION 1: AGRICULTURAL SERVICES/SERVICES AGRICOLES

021 Agricultural Services, Misc Services agricoles, divers

DIVISION 2: FORESTRY/L'EXPLOITATION FORESTIERE

DIVISION 3: FISHING/PECHE

DIVISION 4: MINING/INDUSTRIES DES MINES

052 Gold Mining/Crushing Extraction d'or

DIVISION 5: MANUFACTURING INDUSTRIES/INDUSTRIES MANUFACTURIERES

Major Group/Grand groupe 5.01:

Food and Beverage Industries/Industries des aliments et boissons

101-P	Pork Curing/Packing	Salaison de porc
102	Fish Curing	Salaison de poisson
104	<i>Cheese Factories</i>	<i>Fromageries</i>
105	<i>Flour/Grist Mill</i>	<i>Farine, Moulin à</i>
107	<i>Bakeries</i>	<i>Boulangeries</i>
108-C	Confectionery	Confiserie
108-R	Sugar Refinery	Raffinerie de sucre
109-B	Brewery	Brasserie
109-C	Cider	Cidre
109-D	Distilleries	Distilleries
109-S	Aerated Water	Eaux gazeuses
109-W	Wine	Vin

Major Group/Grand groupe 5.02

Tobacco Products Industries/Industries du tabac

153 Tobacco Working Tabac, Manufacture de

Major Group/Grand groupe 5.03

Rubber Industries/ Industries des produits en caoutchouc

Major Group/Grand groupe 5.04**Leather Industries/Industries du cuir et des produits connexes**

172	<i>Tanneries</i>	<i>Tanneries</i>
174	Boots/Shoes	Bottes/Souliers
179-S	<i>Saddle/Harness Making</i>	<i>Sellerie</i>
179-T	<i>Trunks/Boxes</i>	<i>Valises/Boîtes</i>

Major Group/Grand groupe 5.05**Textile Industries/Industries textiles**

182	<i>Woolen Cloth Factories</i>	<i>Drap, Fabriques de</i>
182-S	Spinning Wool	Filage de laine
182-W	Weaver	Tisserand
184	<i>Rope/Twine Making</i>	<i>Corderies</i>
187	Sail Loft	Voilerie
189	Textiles, Misc	Textiles divers
189-D	<i>Dyeing/Scouring</i>	<i>Teintureries/Dégraissage</i>
189-F	<i>Scutching Mills (flax)</i>	<i>Broyer le lin, Moulins à</i>
189-W	<i>Carding/Fulling Mills</i>	<i>Carder/fouler, Moulins à</i>

Major Group/Grand groupe 5.06**Knitting Mills/Bonneterie**

239	<i>Hosiery</i>	<i>Bonneterie</i>
239-K	Stockings/Mittens/Hand Knitter	Bas/Mitaines/Tricoteur

Major Group/Grand groupe 5.07**Clothing Industries/Industries de l'habillement**

242	Clothing	Habillements
243	Tailors/Clothiers	Habillements pour hommes
245	Children's Clothing	Vêtements pour enfants
246	<i>Furrier/Hatter</i>	<i>Pelleterie/Chapellerie</i>
249	Clothing, miscellaneous	Vêtements divers
249-H	Hats (except fur)	Chapeaux
249-M	Milliner	Modiste

Major Group/Grand groupe 5.08**Wood Industries/Industries du bois**

251	<i>Saw Mill</i>	<i>Moulin à scie</i>
-----	-----------------	----------------------

251-S	<i>Shingle Making</i>	<i>Bardeaux, Confection de</i>
254	<i>Sashes/Doors/Blinds</i>	<i>Portes/Fenêtres</i>
256	<i>Basket Making</i>	<i>Vanneries</i>
258	<i>Coffins</i>	<i>Cercueils</i>
259	<i>Wood Products, Misc</i>	<i>Produits de bois, divers</i>
259-C	<i>Cooperage</i>	<i>Tonnellerie</i>
259-G	<i>Carving/Gilding</i>	<i>Sculpture/Dorure</i>
259-W	<i>Wood Turning Establishments</i>	<i>Tours à bois</i>

Major Group/Grand groupe 5.09
Furniture Industries/Industries du meuble

261	<i>Cabinets/Furniture</i>	<i>Meublerie</i>
266	<i>Furniture, Misc, incl upholstery</i>	<i>Meubles divers</i>

Major Group/Grand groupe 5.10
Paper Industries/Industries du papier

Major Group/Grand groupe 5.11
Printing and Publishing/Imprimerie et édition

286	<i>Printing office</i>	<i>Imprimerie</i>
287-B	<i>Book Binding</i>	<i>Reliure</i>
289	<i>Printing/Publishing incl newspapers</i>	<i>Imprimerie/Publication de journaux</i>

Major Group/Grand groupe 5.12
Primary Metal Industries/Première transformation des métaux

294	<i>Foundry/Castings</i>	<i>Fonderie/Fonte</i>
-----	-------------------------	-----------------------

Major Group/Grand groupe 5.13
Metal Fabricating Industries/Fabrication de produits en métal

304-T	<i>Tinsmith</i>	<i>Ferblantier</i>
306	<i>Agricultural Hand Tools</i>	<i>Outils agricoles manuels</i>
306-T	<i>Axes</i>	<i>Haches</i>
307-S	<i>Stoves</i>	<i>Poêles</i>
309	<i>Metal Fabricating, Misc</i>	<i>Fabrication de métaux, divers</i>

Major Group/Grand groupe 5.14**Machinery Industries/Industries de la machinerie**

311	<i>Agricultural Implements</i>	<i>Instruments aratoires</i>
315	<i>Machinery, Misc</i>	<i>Machinerie/Équipement divers</i>
315-H	<i>Spinning Wheels</i>	<i>Rouets</i>
315-P	<i>Pumps</i>	<i>Pompes</i>

Major Group/Grand groupe 5.15**Transportation Equipment Industries**

326	<i>Locomotives</i>	<i>Locomotives</i>
326	<i>Railway Cars</i>	<i>Chars pour chemin de fer</i>
327	<i>Ship Yards</i>	<i>Navires, Construction de</i>
327-M	<i>Ship Material</i>	<i>Appareux de navire</i>
328	<i>Boat Building</i>	<i>Chaloupes, Construction de</i>
329	<i>Carriage Making</i>	<i>Carrosserie</i>
329-P	<i>Wheels (* and other parts)</i>	<i>Roues</i>

Major Group/Grand groupe 5.17**Non-Metallic Mineral Products/Industries des produits minéraux non métalliques**

351-B	<i>Brick/Tile Making</i>	<i>Briqueterie/Tuilerie</i>
351-P	<i>Pottery</i>	<i>Poterie</i>
353	<i>Stone/Marble Establishments</i>	<i>Marbrerie</i>
357	<i>Grindstone Works</i>	<i>Meules à aiguiser, Fabrique de</i>
358-L	<i>Lime Kilns</i>	<i>Fours à chaux</i>
359-P	<i>Gypsum Mills</i>	<i>Plâtre, Moulin à</i>

Major Group/Grand groupe 5.18**Petroleum and Coal Products Industries/Fabrication de produits du pétrole et du charbon**

365	<i>Oil Refineries</i>	<i>Épuration de l'huile</i>
-----	-----------------------	-----------------------------

Major Group/Grand groupe 5.19**Chemical Products Industries/Industries chimiques**

374	<i>Drugs</i>	<i>Drogues</i>
376	<i>Soap/Candle Making</i>	<i>Savonneries/Chandelleries</i>
378-P	<i>Pot/Pearl Asheries</i>	<i>Potasseries</i>

379-H	<i>Charcoal Burning</i>	<i>Charbonnerie</i>
379-M	<i>Matches</i>	<i>Allumettes</i>

Major Group/Grand groupe 5.20**Miscellaneous Manufacturing Industries/Autres industries manufacturières**

391	<i>Scientific/Professional Equipment</i>	<i>Équipement scientifique</i>
392	<i>Gold/Silver Smithing</i>	<i>Orfèvrerie</i>
392	<i>Jeweler/Watchmaker</i>	<i>Bijoutier/Horloger</i>
399	<i>Miscellaneous wares</i>	<i>Articles variés</i>
399-B	<i>Brooms/Brushes</i>	<i>Brosserie</i>
399-P	<i>Musical Instruments</i>	<i>Instruments de musique</i>

DIVISION 6: CONSTRUCTION INDUSTRY/BATIMENT ET TRAVAUX PUBLICS

421-C	<i>Carpenter/Joiner (buildings)</i>	<i>Charpentier/Menuisier (bâtiments)</i>
421-D	<i>Painter/Glazier</i>	<i>Peintre en bâtiments</i>
421-P	<i>Plumber/Gasfitter</i>	<i>Plombier/Gazier</i>
421-S	<i>Stonemason/Bricklayer</i>	<i>Maçon</i>

DIVISION 7: UTILITIES/AUTRES SERVICES PUBLICS

574	<i>Gas Works</i>	<i>Gaz, fabrique de</i>
-----	------------------	-------------------------

DIVISION 8: TRADE/COMMERCE DE GROS

658	<i>Carriage Repairs</i>	<i>Réparations de voitures</i>
695	<i>Jeweler/Watchmaker Repairs</i>	<i>Bijoutier/Horloger: réparations</i>

DIVISION 10: BUSINESS AND PERSONAL SERVICES/SERVICES PERSONNELS

896	<i>Blacksmithing</i>	<i>Forges</i>
-----	----------------------	---------------

NOTE: *Industry types printed in italics are the terms used in the published tabulations of the 1871 Census reports.*

Appendix A-3: CANIND71 database: alphabetical list of variables

AVWAGE:	Average monthly wage per employee in dollars and cents -- a derived variable calculated only when data for TOTEMP, WAGES and MONTH were not missing.
CDID:	Census District code, comprising initial letter for province followed by sequential three-digit number, from 001 in southwestern Ontario to 206 in northeastern Nova Scotia.
CDISTRIC:	Census District name, with qualifiers such as North, South, East, and West placed after the main name.
CED:	Census enumerator's division, a letter and number combination for all or, more often, part of a CSD.
COMMENTS:	Remarks added on the schedule, providing additional information about the establishment (up to 200 characters).
CSD:	Census Sub-District name, usually corresponding to basic municipal unit such as township, town or village in Ontario, and also to ward of major cities. Qualifiers such as North, etc placed after the main name and abbreviated to N, etc. Urban municipal status indicated by T for Town/Ville and V for Village.
EMPBOY:	Boys under 16 years (number employed).
EMPGIRL:	Girls under 16 years (number employed).
EMPMEN:	Men over 16 years (number employed).
EMPWOM:	Women over 16 years (number employed).
FIXCAP:	Fixed capital invested, in dollars.
FLOCAP:	Floating capital employed, in dollars.
FORCE:	Nominal force of moving power, stated in horse power units.
MONTH:	Number of working months in the year.
PQUANT1:	Quantity of first product, with provision for up to 12 products per establishment.
PROD1:	Name of first product, with provision for up to 12 products per establishment.
PROP:	Has an "F" entered if the proprietor had a woman's name.
PROPRIOR:	Name of proprietor or company.

PUNIT1: Unit of measurement of first product, according to abbreviations in Appendix A-4, with provision for up to 12 products per establishment.

PVALUE1: Value of first product, in dollars, with provision for up to 12 products per establishment.

RAWMAT1: Name of first raw material, with provision for up to 12 raw materials per establishment.

REFNUM: Unique number of individual establishment in the database.

RQUANT1: Quantity of first raw material, with provision for up to 12 raw materials per establishment.

RUNIT1: Unit of measurement of first raw material, according to abbreviations in Appendix A-4 , with provision for up to 12 raw materials per establishment.

RVALUE1: Value of first raw material, in dollars, with provision for up to 12 raw materials per establishment.

SEC: Major industry group or sector, in which basic SIC codes are grouped together as in Table 1. Derived from SIC variable following.

SIC: Standard Industrial Classification code, to accommodate variant suffixes and composites, as in Appendix A-2.

SUMPROC: Total value of products, the sum of PVALUE1 + PVALUE2 ... etc, in dollars, a derived variable.

SUMRAWC: Total value of raw materials, the sum of RVALUE1 + RVALUE2 ... etc, in dollars, a derived variable.

TOTEMP: Sum of all employees, EMPMEN, EMPWOM, EMPBOY and EMPGIRL, a derived variable.

TYPEEST: Type of industrial establishment, in the natural language of the manuscript schedules.

TYPEPOW: Type of moving power other than manual, with a controlled vocabulary of types: Water, Steam, W/S (Water/Steam), Horse, Wind.

VADD: Value added in manufacturing in dollars, a variable derived by subtracting SUMRAWC from SUMPROC when neither is missing.

WAGES: Aggregate amount of yearly wages, in dollars.

Appendix A-4

ENGLISH UNITS OF MEASUREMENT AND ABBREVIATIONS, 1871 CENSUS

ENGLISH UNIT	ABBREV	FRENCH EQUIVALENT
1.25 METRES		AUNE/AULNE (F)
1.5 ENGLISH ACRES	A	ARPENT (M)
ACRE	A	ACRE (F)
BALE		BALLE (F)
BOARD FEET	B FT	PIED DE PLANCHE
BOARD MEASURE	BM	MESURE DE PLANCHE
BOARD MEASURE		PLANCHE MESURE
BOARD/PLANK		PLANCHE (F)
BOTTLE		BOUTEILLE (M)
BOX/CAN		BOITE (F)
BUNCH/TRUSS/BALE		BOTTE (F)
BUNDLE (OF PAPER)		LIASSE (F) (DE PAPIER)
BUNDLE (OF WOOD)		FAGOT (M) (DE BOIS)
BUSHEL	BU	BOISSEAU (M)
BUTTER TUB		TINETTE (F)
CASE/BOX/CHEST		CAISSE (F)
CASK/BARREL	BBL	TONNEAU (M)
CAULDRON	CLDN	CHAUDRON (M)
CENSUS STANDARD LOG	CSL	BILLOT ?
CORD	CD	CORDE (F)
CUBIC FEET	CU FT	PIED CUBIQUE
CUBIC YARD	CU YD	VERGE CUBIQUE
DOZEN	DOZ	DOUZAIN (F)
DOZENS OF BOTTLES	DOZ B	DOUZAINES DE BOUTEILLES
FEET	FT	PIED (M)
FEET BOARD MEASURE	FT BM	PIED DE MESURE DE PLANCHE
FOOT MEASURE		PIED MESURE
GALLON	GAL	GALLON (M)
GARMENT	GRMENT	VETEMENT (M)
GROSS		GROSSE (F)
HUNDREDWEIGHT	CWT	QUINTAL
INCH		POUCE (M)
KEG		PETIT TONNEAU (M)
KEG/CASK/BARREL, SMALL	BBL	BARIL (M)
LARGE BARREL/HOGSHEAD	BBL	BARRIQUE (F)
LEGAL UNIT OF MEASURE		ETALON (M)
LOAF		PAIN (M)
LOG		BILLOT (M)
NUMBER	NO	NOMBRE (M)
OUNCE	OZ	ONCE (F)
PACKAGE	PCKG	PAQUET (M)
PAIR	PR	PAIRE (F)
PECK		MINOT (M)
PHIAL		FIOLE (F)
PIECE		PIECE
PIECE		MORCEAU (M)
PLANK FOOR MEASURE		PIED DE MESURE DE MADRIER
PLANK, THICK		MADRIER (M)
POUCH/BAG		POCHE (F)
POUND	LB	LIVRE (F)
PUNCHEON	PNCHN	POINCON (M)
QUART		QUART
REAM (OF PAPER)		RAME (F)

ENGLISH UNITS OF MEASUREMENT AND ABBREVIATIONS, 1871 CENSUS

ENGLISH UNIT	ABBREV	FRENCH EQUIVALENT
ROLL		ROULEAU (M)
SET (OF HARNESSSES)		MENOIRE (F)
SHEAF/BUNDLE		GERBE (F)
SIDE		COTE (M)
SKEIN		ECHEVEAU (M)
SKIN/HIDE		PEAU (F)
SQUARE FEET	SQ FT	PIED CARRE
SQUARE YARD	SQ YD	VERGE CARRE
SUIT		COMPLET (M)
SUPERFICIAL FEET	SQ FT	PIED DE SUPERFICIE
TOISE		TOISE (F)
TON		TONNE (F)
TONNAGE	TONN	TONNAGE (M)
WEB		SANGLE (F)
YARD	YD	VERGE (F)

FRENCH UNITS OF MEASUREMENT AND ABBREVIATIONS, 1871 CENSUS

FRENCH TERM	ABBREV	ENGLISH EQUIVALENT
ACRE (F)	A	ACRE
ARPENT (M)	A	1.5 ENGLISH ACRES
AUNE/AULNE (F)		1.25 METRES
BALLE (F)		BALE
BARIL (M)	BBL	KEG/CASK/BARREL, SMALL
BARRIQUE (F)	BARRIQ	LARGE BARREL/HOGSHEAD
BILLOT (M)		LOG
BILLOT ?	?	CENSUS STANDARD LOG
BOISSEAU (M)	BOISSE	BUSHEL
BOITE (F)		BOX/CAN
BOTTE (F)		BUNCH/TRUSS/BALE
BOUTEILLE (M)		BOTTLE
CAISSE (F)		CASE/BOX/CHEST
CHAUDRON (M)	CHDN	CAULDRON
COMPLET (M)		SUIT
CORDE (F)	CD	CORD
COTE (M)		SIDE
DOUZAINE (F)	DOUZ	DOZEN
DOUZAINES DE BOUTEILLES	DOUZ B	DOZENS OF BOTTLES
ECHEVEAU (M)	ECHEVE	SKEIN
ETALON (M)		LEGAL UNIT OF MEASURE
FAGOT (M) (DE BOIS)		BUNDLE (OF WOOD)
FIOLE (F)		PHIAL
GALLON (M)	GAL	GALLON
GERBE (F)		SHEAF/BUNDLE
GROSSE (F)		GROSS

Appendix A-4 FRENCH UNITS OF MEASUREMENT AND ABBREVIATIONS, 1871 CENSUS

FRENCH TERM	ABBREV	ENGLISH EQUIVALENT
LIASSE (F) (DE PAPIER)		BUNDLE (OF PAPER)
LIVRE (F)		POUND
MADRIER (M)		PLANK, THICK
MENOIRE (F)		SET (OF HARNESSSES)
MESURE DE PLANCHE	MP	BOARD MEASURE
MINOT (M)		PECK
MORCEAU (M)	MORCEA	PIECE
NOMBRE (M)	NO	NUMBER
ONCE (F)		OUNCE
PAIN (M)		LOAF
PAIRE (F)	PR	PAIR
PAQUET (M)		PACKAGE
PEAU (F)		SKIN/HIDE
PETIT TONNEAU (M)	P TONN	KEG
PIECE		PIECE
PIED (M)	PI	FEET
PIED CARRE	PI CA	SQUARE FEET
PIED CUBIQUE	PI CU	CUBIC FEET
PIED DE MESURE DE MADRIER	PI MM	PLANK FOOR MEASURE
PIED DE MESURE DE PLANCHE	PI MP	FEET BOARD MEASURE
PIED DE PLANCHE	PI PL	BOARD FEET
PIED DE SUPERFICIE	PI S	SUPERFICIAL FEET
PIED MESURE	PI M	FOOT MEASURE
PLANCHE (F)	PL	BOARD/PLANK
PLANCHE MESURE	PM	BOARD MEASURE
POCHE (F)		POUCH/BAG
POINCON (M)		PUNCHEON
POUCE (M)		INCH
QUART		QUART
QUINTAL	Q	HUNDREDWEIGHT
RAME (F)		REAM (OF PAPER)
ROULEAU (M)		ROLL
SANGLE (F)		WEB
TINETTE (F)	TINETT	BUTTER TUB
TOISE (F)		TOISE
TONNAGE (M)	TONN	TONNAGE
TONNE (F)		TON
TONNEAU (M)	TONNEA	CASK/BARREL
VERGE (F)		YARD
VERGE CARRE	V CA	SQUARE YARD
VERGE CUBIQUE	V CU	CUBIC YARD
VETEMENT (M)	VTMENT	GARMENT

Name		Address		City		State		Zip	
John Doe		123 Main St		New York		NY		10001	
Jane Smith		456 Elm St		Los Angeles		CA		90001	
Bob Johnson		789 Oak St		Chicago		IL		60601	
Alice Brown		101 Pine St		Houston		TX		77001	
David Wilson		202 Maple St		Phoenix		AZ		85001	
Emily Davis		303 Cedar St		Philadelphia		PA		19101	
Frank Miller		404 Birch St		San Antonio		TX		78101	
Grace Lee		505 Spruce St		San Diego		CA		92101	
Henry White		606 Willow St		Dallas		TX		75201	
Ivy Green		707 Ash St		San Jose		CA		95101	
Jack Black		808 Hickory St		Austin		TX		78701	
Karen Blue		909 Sycamore St		Jacksonville		FL		32201	
Leo Red		1010 Magnolia St		Fort Worth		TX		76101	
Mia Purple		1111 Dogwood St		Columbus		OH		43201	
Noah Gold		1212 Redwood St		Indianapolis		IN		46201	
Olivia Silver		1313 Cypress St		San Francisco		CA		94101	
Peter Bronze		1414 Juniper St		Nashville		TN		37201	
Quinn Copper		1515 Fir St		Portland		OR		97201	
Ryan Iron		1616 Hemlock St		Memphis		TN		38101	
Sara Steel		1717 Spruce St		San Jose		CA		95101	
Tina Lead		1818 Cedar St		San Antonio		TX		78101	
Uma Zinc		1919 Birch St		San Diego		CA		92101	
Victor Tin		2020 Oak St		Dallas		TX		75201	
Wendy Nickel		2121 Pine St		San Jose		CA		95101	
Xavier Cobalt		2222 Elm St		Austin		TX		78701	
Yara Manganese		2323 Maple St		Jacksonville		FL		32201	
Zoe Vanadium		2424 Spruce St		Fort Worth		TX		76101	

University of Waterloo
Library

University of Waterloo
Library

University of Waterloo
Library

